Vol. 42, Parts 1 and 2 June 1955

BIOMETRIKA

FOUNDED BY
W. F. R. WELDON, FRANCIS GALTON AND KARL PEARSON

MANAGING EDITOR
E. S. PEARSON

ASSOCIATE EDITORS
M. G. KENDALL JOHN WISHART

in consultation with

HARALD CRAMÉR J. B. S. HALDANE
F. N. DAVID H. O. HARTLEY
R. C. GEARY D. G. KENDALL

Re-issued photo-litho 1966

ISSUED BY
THE BIOMETRIKA OFFICE, UNIVERSITY COLLEGE, LONDON
PRINTED AT THE UNIVERSITY PRESS, CAMBRIDGE

[*Issued 3 June 1955*]

VOLUME 42, PARTS 1 AND 2 JUNE 1955

STUDIES IN THE HISTORY OF PROBABILITY AND STATISTICS*

I. DICING AND GAMING (A NOTE ON THE HISTORY OF PROBABILITY)

BY F. N. DAVID

University College, London

'See, this is new? It hath been already of old time.' (Ecclesiastes i. 10.)

1. A cynical archaeologist remarked recently that a symptom of decadence in a civilization is when men become interested in their own history, and he added that in the unlikely eventuality of any proof being required of the decadence of this phase of *Homo sapiens* it could be found in the present-day interest in archaeology. Most generalizations of a sweeping character such as this are unacceptable, chiefly because there is no way of putting them to proof; but the present interest of scientists in general, and of statisticians in particular, in the origins of scientific thought, is far from implying the decadence of science, whatever may be implied by an interest in the arts.

It is inviting, and at the same time profitless, to speculate why modern scientists have such an interest. The possibility of deciding priority of discovery which concerned the Victorian scientist so closely does not cause much controversy to-day, for the modern scientist would hold that to ascribe any discovery in the field of science to any single person is unrealistic. Thus, while we are taught at school, for example, that Newton and Leibnitz separately and independently 'discovered' the differential calculus, it would perhaps be more appropriate to say that Newton and Leibnitz each supplied the last link in the chain of reasoning which gave us the differential calculus—a chain which can be traced back through Pierre Fermat, Barrow, Torricelli and Galileo, and that it is surprising that there were only two mathematicians who did this.

Mathematics is essentially an expression of thought in which we build on the mental effort of our forerunners, and probability is no exception to this general rule. The real difficulty we meet with in trying to trace probability back to its origins is that it started essentially as an empirical science and developed only lately on the mathematical side. It is hard to say where in time the change came from empiricism to mathematical formalism as it appears to have taken place over hundreds of years; and the claims put forward for Pascal and Fermat as the creators of probability theory cannot entirely be substantiated.

2. When man first started to play games of chance is a time problem we shall never clearly resolve. We may place on record that it is a commonplace thing for archaeologists to find a preponderance of astragali† among the bones of animals dug up on prehistoric sites. One archaeologist stated that he had found up to seven times as many as any other bone, another put the figure at 500 (sic!), while yet a third, refusing to be drawn to a figure, stated that they were many. This fact has probably little significance. The astragalus has little marrow in it and was possibly not worth cracking for the sake of its contents as were the long bones; it is knobbly and presents no flat curves for drawing as does the shoulder

* [*Editorial note.* It is hoped to publish articles by a number of different authors under this general heading.]

† The *astragalus* is a small bone in the ankle, immediately under the *talus* or heel-bone. See Pl. 2*a*.

UNIVERSITY OF UTAH LIBRARIES

blade for example. All we may do is to place on record that round about 40,000 years ago there were large numbers of the astragali of sheep, goat and deer lying about.

The astragali of animals with hooves are different from those with feet such as man, dog and cat. From the comparison in Text-fig. 1 we note how in the case of the dog the astragalus is developed on one side to allow for the support of the bones of the feet. The astragalus of the hooved animal is almost symmetrical about a longitudinal axis and it is a pleasant toy to play with. In France and Greece children still play games with them in the streets, and it is possible to buy pieces of metal fashioned into an idealized shape but still recognizable as astragali.

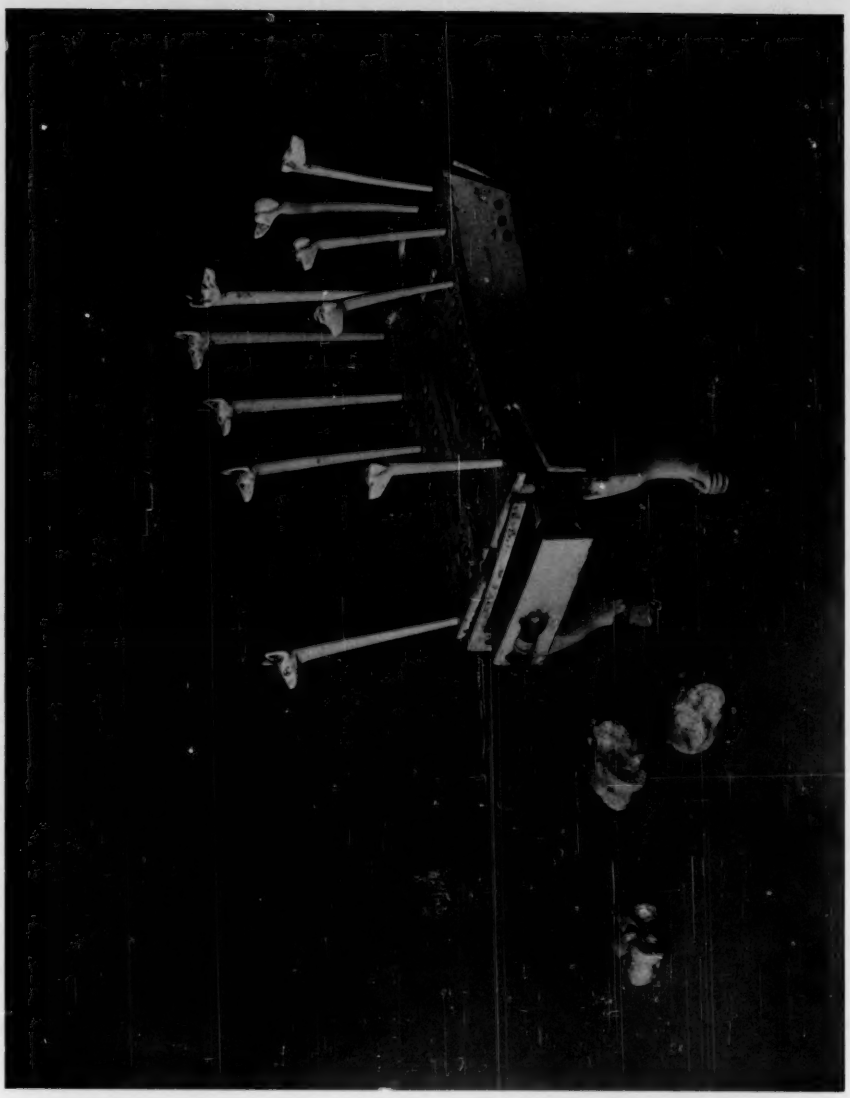
Text-fig. 1. Drawings of the astragalus in sheep and dog, natural size.

3. Some time between prehistoric man of four hundred centuries ago and the beginning of the third millennium before Christ *Homo sapiens* invented games and among these games, games of chance. We know from paintings, terra-cotta groups, etc., that the astragalus was used in Greece like the ancient quoit,* but there is no doubt from paintings on tombs in Egypt and excavated material that the use of the astragalus in games where it is desired to move counters was well established by the time of the First Dynasty. In one painting a nobleman, shown playing a game in his after-life, delicately poises an astragalus on his finger tip, a board with 'men' in front of him. A typical game of *c.* 1800 B.C. is that of 'Hounds and Jackals' illustrated in Pl. 1. The game seems similar to our present-day 'Snakes and Ladders'; the hounds and jackals were moved according to some rule by throwing the astragali found with the game and shown in the figure. Variants of this game were undoubtedly played from the time of the First Dynasty (*c.* 3500 B.C.).

It is possible but not altogether likely that these games originated in Egypt. They certainly did not originate in Greece, as has been claimed for reasons which we shall give later. However, Herodotus, the first Greek historian, like his present-day counterparts, was willing to believe that the Greeks (or allied peoples) had invented nearly everything. His claim that the Lydians introduced coinage has about as much foundation as his claim regarding games of chance. He writes (*c.* 500 B.C.) about the famine in Lydia (which was *c.* 1500 B.C.) as follows:

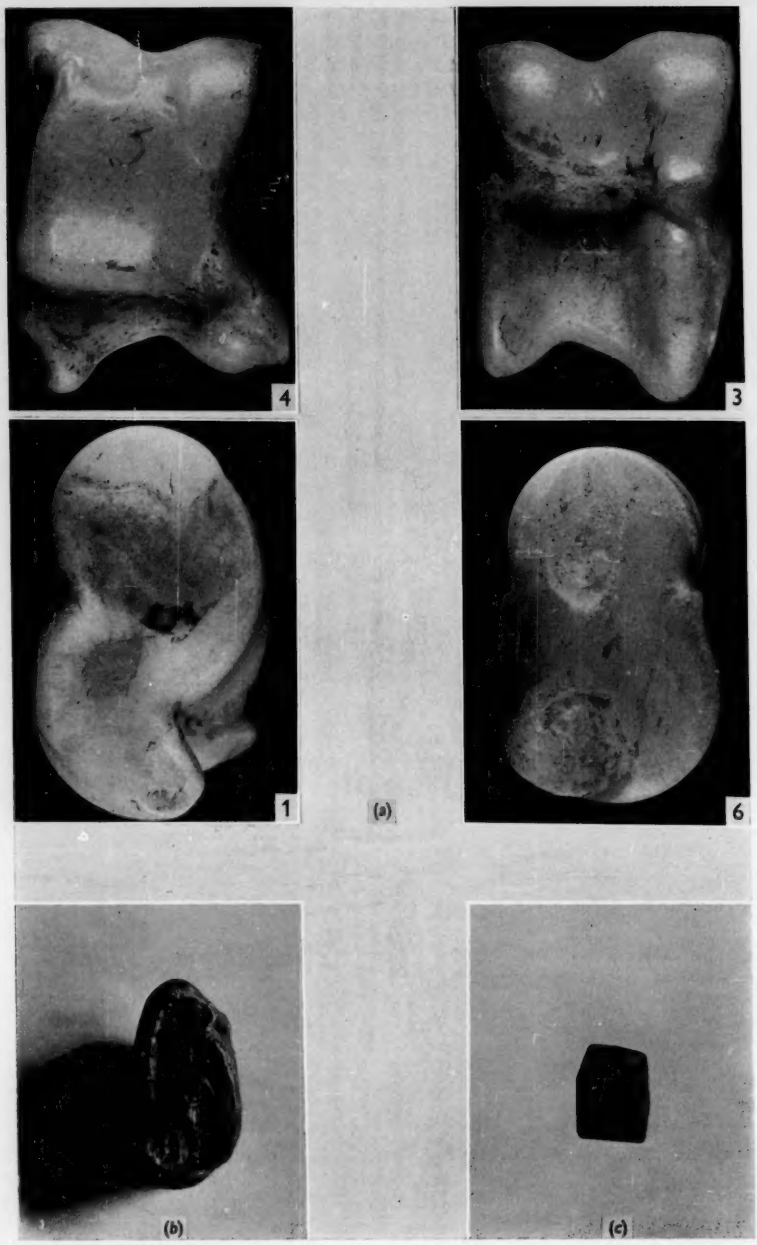
> The Lydians have very nearly the same customs as the Greeks. They were the first nation to introduce the use of gold and silver coins and the first to sell goods by retail. They claim also the invention of all games which are common to them with the Greeks. These they declare that they invented about the time that they colonized Tyrrhenia, an event of which they give the following account. In the days of Atys, the son of Manes, there was great scarcity through the whole land of Lydia. For some time the Lydians bore the affliction patiently, but finding that it did not pass away,

* From the name 'knucklebone' we might infer that among the early games were those in which the astragali were balanced on the bones of the knuckles and then tossed and caught again.



The game of Hounds and Jackals.

(*Facing p.* 2)


4
3
1
(a)
6
(b)
(c)

they set to work to devise remedies for the evil. Various expedients were discovered by various persons; dice and huckle-bones (i.e. astragali) and ball and all such games were invented, except tables,* the invention of which they do not claim as theirs. The plan adopted against famine was to engage in games on one day so entirely as not to feel any craving for food, and the next day to eat and abstain from games. In this way they passed eighteen years.

In yet another commentary we are told that games of chance were invented during the Trojan war by Palamedes. During the 10-year investment of the city of Troy various games were invented to prevent the soldiers' morale suffering from boredom.

4. The game of ball is mentioned by Homer and according to Plato was evolved in Egypt. It is not, however, a game of chance. The story of dice we shall return to, but we may first carry the story of the astragalus a little further. In the early part of the first millennium it would seem that astragali were used by both adults and children for their leisure games. Homer (*c.* 900 B.C.) tells us that when Patroclus was a small boy he became so angry with his opponent while playing a game of knucklebones that he nearly killed him. Another writer of the same period tells us that students played knucklebones everywhere, that they were acclaimed as presents and that as a prize for handwriting one student was given eighty astragali all at once! It is not difficult to imagine the small boys of that era collecting astragali as they collected marbles, much as the boys of our own era still do.

That the astragalus was used commonly in the gaming which the Greeks and later the Romans conducted with zeal and passion, the references in the literature of that time leave no room for doubt. One of the chief games may have been the simple one of throwing four astragali together and noting which sides fell uppermost. The astragalus has only four sides on which it may rest, and it has been deduced, among others by Nicolas Leonicus Thomeus (1456–1531), that a common method of enumeration was that the upper side, broad and slightly convex counted 4, the lower side broad and hollowed 3, the lateral side narrow and flat 1 and the other lateral which is slightly hollow 6. These aspects of a sheep's astragalus are shown in Pl. 2*a*. (With present-day astragali the probabilities of scoring 1 and 6 are each approximately equal to 1/10 and those of 3 and 4 approximately 4/10.) The worst throw for the Greeks with one bone was unity which they called the dog, and sometimes the vulture. The best of all throws with four knucklebones was the throw of Venus when all four sides were different which has an actual probability of about 1/26. But at different times and in different games the numbers must have been varied, for the throw of Euripides with four astragali, discussed by several fifteenth-century writers, was worth 40. How the bones fell to achieve this result is not stated, although Cardano writing in the sixteenth-century states that it was four fours. (Probability *c.* 1/39.)

5. In classical Rome the astragalus was imitated in carved stone with figures and scenes incised on the sides. A typical example is illustrated in Pl. 2*b*. Stone astragali have also been found in Egypt. At this time too we have the production of lewd figures in metal or bone varying in size from about 1 cm. to over 1 in. in height. That these figures were used for gaming may be deduced from the fact that the six possible positions in which the figure may fall are each marked with a number of dots.†

Besides the astragali it appears possible that throwing sticks were also used for games of chance, although it may be that they had a greater religious significance; we shall return

* This may have been an early form of backgammon or may have been shuffle-board.

† I have not tested these figures for bias. They are a development, I think, of dice rather than astragali.

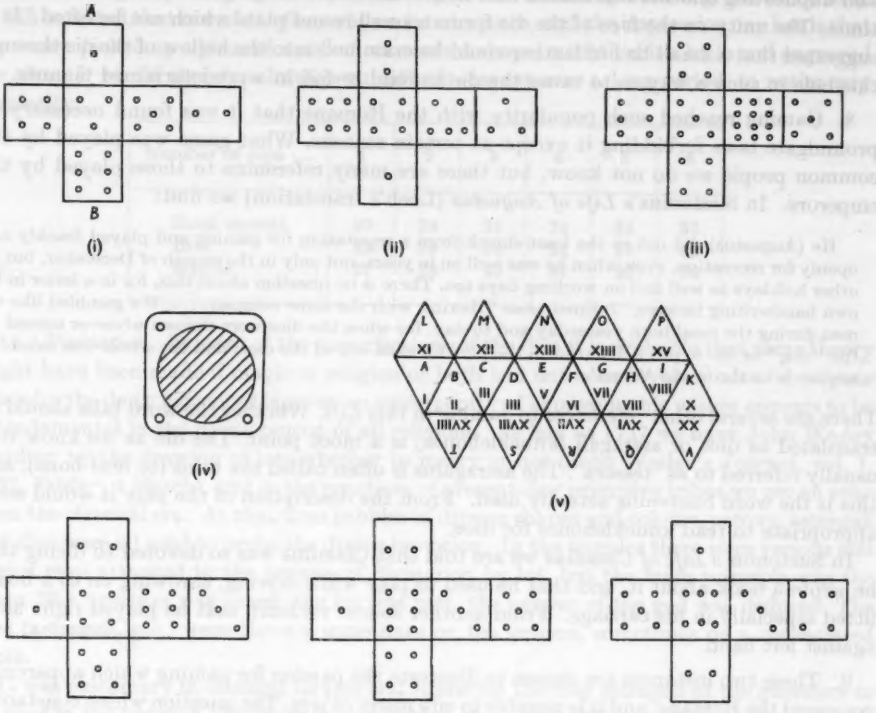
to this point later. The throwing stick was made of wood or ivory and was often approximately 3 in. in length with cross-section when square of about 1 cm. each way. Such throwing sticks were known to the ancient Britons, to the Greeks, the Romans, the Egyptians and to the Maya Indians of the American continent. Sometimes the sticks are elliptical in cross-section with major axis of approximately 1 cm., but they are all alike in having only four numbers on them, one at each end of the upper face and one at each end of the lower parallel face. In the European throwing sticks the majority of numbers are marked by small engraved circles, but they are sometimes indicated by cuts in the wood or ivory. The Maya throwing sticks are marked by coloured scratches in ivory. The actual numbers marked vary. They are mostly 1, 2, 5, 6, but 3 and 4 have also been noticed. These throwing sticks are of little importance in gaming. They are mentioned because it is interesting to note the likelihood that gaming originated at many points, and, although this is a remark one could not defend, that it possibly was originally a debasement of a religious ceremony.

6. The six-sided die may have been obtained from the astragalus by grinding it down until it formed a rough cube. The Musée de Louvre has several astragali which have been treated in this way, but one cannot imagine they formed very satisfactory dice. The honeycombed (or cancellous) bone tissue has been exposed in several places and the crude die would clearly not have a long life. Whether the die was evolved in this way or not the evolution must have taken place some considerable time before Christ. The earliest die known was excavated in northern Iraq and is dated at the beginning of the third millennium. It is described as being of well-fired buff pottery. The dots are arranged as shown in Text-fig. 2(i), the edges at A and B being imagined folded away from the reader. It will be noted that the opposite points are in consecutive order, 2 opposite 3, 4 opposite 5 and 6 opposite 1.

A die excavated in Mohenjo-Daro (Ancient India) is also dated as third millennium, and it is also made of hard buff pottery. The order of the points is again consecutive, but this time we have 1 opposite 2, 3 opposite 4, 5 opposite 6. Few other dice have been recorded in this millennium. At the time of the XVIIIth dynasty in Egypt (*c.* 1400 B.C.) a die with the markings shown in Text-fig. 2(ii) must have been in play. The arrangement of the five dots is unusual. Somewhere about this time, however, the arrangement of the numbers settled down to the familiar two-partitions of 7 opposite one another as shown in Text-fig. 2(vi), which arrangement has persisted until to-day. Out of records, collected by the present writer, of some fifty dice of classical times made of crystal, ivory, sandstone, ironstone, wood and other materials, forty had the two-partition of 7 arrangement. A twelfth-century (A.D.) Greek bishop wrote that this was the way in which a die should be marked, and a sixteenth-century gambler theorizes that this arrangement was chosen to make it easy to check whether all the numbers had been marked on the die and no figure duplicated at the expense of leaving out another. One die of the first millennium is said to have 9 opposite 6, 5 opposite 3, and 4 opposite 2. It may have been especially made for a particular game; alternatively, it is possible that it may have some ceremonial significance. This is possibly also true of a die marked as in Text-fig. 2(iii), although it might have been a die used for cheating.

7. That dice were used in Egypt is clear from the XVIIIth-dynasty specimen. It is thought, however, that dicing did not become common until the advent of the Ptolemaic dynasty (300 to 30 B.C.) which originated from Greece. Several dice are known of this

period including a beautiful specimen in hard brown limestone of side *c.* 1 in., which has the sacred symbols of Osiris, Horus, Isis, Nebhat, Hathor and Horhudet engraved on the six sides. This would almost undoubtedly have been used for some form of divination rites.



Text-fig. 2

Dice in Britain were in a very primitive state at the time of Christ. The pieces used then were formed by roughly squaring the long bone of an animal and cutting it into sections to form objects approximately cubical in shape. The marrow was taken out leaving a hollow square cylinder of which the cross-section in diagrammatic form is sketched in Text-fig. 2 (iv). (The British Museum has two of these.) These primitive dice had the two partition of 7 arrangement, 3 being opposite 4 on the hollow ends. Some dice had a 3 on each end and no 4. Several dice of this kind have been excavated in the chalk and flint country and dated late in the first millennium.

The working out of the geometry of solid figures by Greek mathematicians appears to have been followed almost immediately by the construction of polyhedral dice. A beautiful icosahedron in rock crystal now at the Musée de Louvre is the most famous of these. (In the diagram of Text-fig. 2 (v) it may be imagined the outline is folded away from the reader.) A figure with 19 faces badly cut but apparently imagined to be rectangular has a roman digit on each face from I to X, and above that the numbers rise by tens to C. The

number LXXX is missing, but the number XX appears twice on one face. There was also a die of 18 faces probably formed by beating out a cubical die, a die with 14 faces and so on.

Faked dice were also not unknown. Apart from the device of leaving out one number and duplicating another it is stated that hollow dice have been found dating from Roman times. The unity on the face of the die forms a small round plate which can be lifted. It is suggested that a small ball of leather could be crammed into the hollow of the die through this hole in such a way as to cause the die to tend to fall in a predetermined manner.

8. Gaming reached such popularity with the Romans that it was found necessary to promulgate laws forbidding it except at certain seasons. What game was played by the common people we do not know, but there are many references to those played by the emperors. In Suetonius's *Life of Augustus* (Loeb's translation) we find:

> He (Augustus) did not in the least shrink from a reputation for gaming and played frankly and openly for recreation, even when he was well on in years, not only in the month of December, but on other holidays as well and on working days too. There is no question about this, for in a letter in his own handwriting he says, 'I dined, dear Tiberius, with the same company; ... We gambled like old men during the meal both yesterday and to-day, for when the dice were thrown whoever turned up the "dog" or the 6, put a denarius in the pool for each one of the dice, and the whole was taken by anyone who threw the Venus'.

There are several other references to gaming in this *Life*. Whether the word talis should be translated as dice or astragali (knucklebones) is a moot point. The die as we know it is usually referred to as 'tessera'. The astragalus is often called the talus (or heel-bone), and this is the word Suetonius actually used. From the description of the play it would seem appropriate to read knucklebones for dice.

In Suetonius's *Life of Claudius* we are told that Claudius was so devoted to dicing that he wrote a book about it, and that he used to play while driving, throwing on to a board fitted especially in his carriage. From another source we learn that he played right hand against left hand.

9. These two instances are chosen to illustrate the passion for gaming which apparently possessed the Romans, and it is possible to cite many others. The question which constantly recurs to one while studying these games of the past is 'Why did not someone notice the equi-proportionality property of the fall of the die?' It is understandable that no theory was made to describe the fall of the astragalus. But the Greeks had performed the necessary abstraction of thought to make the mathematical idealization of the cube (and other solid figures); at first sight it seems curious that mathematicians did not then go on a little further and give equal weight to each side of the cube and so on. For if dicing and gaming generally were carried on by so many persons for so long that it was thought necessary to prohibit them, surely someone must have noticed that with a cube on the average any one side turned up as frequently as any other? We can only make guesses on this point, but it would seem to the writer that there are two possible explanations, the imperfections of the dice and their use in religious ceremonies.

10. *Imperfect dice.* We speak of a true or a fair die nowadays when we mean that there is no bias apparent when the die is thrown. In Roman times, and presumably earlier, it seems to have been the exception rather than the rule for the die to be true. Many dice of the classical period have been thrown by the writer and they were nearly all biased but not all in the same way. For example, three classical dice from the British Museum gave the results shown in the table from 204 throws each. The arrangement of the pips on the dice

were as in Text-fig. 2(vi), (vii) and (viii). The rock crystal is a beautifully made die; the others are a little primitive, and the sides of the iron die are only approximately parallel. The marble and the iron dice are obviously biased, and this was true of many of the other dice examined. A photograph of a wooden die of the classical period is given in Pl. 2(*c*). It will be noted that one of the faces shown is not square, and the impression one has is that the owner picked up a piece of wood of convenient shape, smoothed it a little and engraved the pips. It would therefore have been difficult, except over a long period, to notice any regularity.

Number of pips ...	1	2	3	4	5	6
Rock crystal	30	38	31	34	34	37
Iron	35	39	30	21	37	42
Marble	27	28	23	47	25	54

11. *Divination.* In spite of the imperfections of the dice it is probable that some theory might have been made if magic or religion or both had not been involved also. A scheme whereby the deity consulted is given an opportunity of expressing his wishes appears to be a fundamental in the development of all religions. As late as 1737 we have John Wesley deciding by the drawing of lots whether to marry or not (John Wesley's *Journal*, vol. 1, 1737, Friday, 4 March), and in the practices of present-day primitive tribes we get an echo from the classical era. At that time pebbles of diverse shapes and colours, arrows, astragali and dice were all used to probe the divine intention. In the temples there were various and varied rites attached to the process of divination by lot, but the main principle was the same. The question was posed, the lot was cast, the answer of the god was deduced. The dice (astragali, etc.) were thrown sometimes on the ground, sometimes on a consecrated table.

It was customary in classical Greece and Rome for the four astragali of the gamblers to be used in the temples. The prediction was that the throw of Venus (1, 3, 4, 6 uppermost) was favourable and the dogs unfavourable. In the temple of the oracles tablets were hung up and the priest, or possibly the suppliant, interpreted the throw of the four bones by reference to the tablets. Cases have been recorded, however, where five astragali were used. Greek inscriptions found in Asia Minor give a fairly complete record of how the throws of five were interpreted. Each throw was given the name of a god. Thus Sir James Frazer translates (commentary on Pausanias):

> 1. 3. 3. 4. 4 = 15 *The throw of Saviour Zeus*
> One one, two threes, two fours,
> The deed which thou meditatest, go, do it boldly.
> Put thy hand to it. The gods have given these favourable omens.
> Shrink not from it in thy mind. For no evil shall befall thee.

It is not clear whether the order of the numbers is important or not. If order does not matter then the probability of this throw is about 0·08.* The tesserae of the gambler were also used in divination ceremonies as well as the astragali, and it is possible that the same interpretation was given to the numbers falling uppermost, although the presence of the 2 and the 5 would make this a little awkward.

* I propose to write at greater length on 'divination probabilities' on a further occasion.

In addition to the divination carried out by the priests it was apparently a commonplace for individuals to perform acts of divination with regard to events in their daily lives. Thus Lucian, telling the story of the young man who fell madly in love with Praxiteles's Venus of Cnidos, writes:

> He threw four knucklebones on to the table and committed his hopes to the throw. If he threw well, particularly if he obtained the image of the goddess herself, no two showing the same number, he adored the goddess, and was in high hopes of gratifying his passion: if he threw badly, as usually happens, and got an unlucky combination, he called down imprecations on all Cnidos, and was as much overcome by grief as if he had suffered some personal loss.

Again we have from Propertius:

> When I was seeking Venus (i.e. good fortune) with favourable tali, the damned dogs always leaped out.

12. It is perhaps of interest here to interpolate a note on divination as reported practised by the Buddhists of present-day Tibet. According to Hastings (*Dictionary of Comparative Religions*) the simplest method is carried out by the people themselves. Many laymen are equipped with a pocket divination manual (*mô-pe*) and the augury found by casting lots. This lot-casting can either be odds and evens (the random pouring of grain, pebbles or coins from a horn, cup, etc.) or dice on a sacred board or cards on which there are magic signs, or sheets or passages of scriptures drawn from a bowl. The reincarnation prediction is, it is said by Waddell,* usually carried out by a priest. The rebirth chart seen by the writer consists of 56 2 in. squares (8×7). Each square corresponds to a future state. A six-sided die with letters on it is thrown down on to the rebirth chart, and according to the square on which it lands and the letter which falls uppermost so the priest predicts. Waddell, who visited Tibet as a member of a British Mission, obtained one of these charts and a die (*c.* 1893). He remarks: 'The dice (sic!) accompanying my board seems to have been loaded so as to show up the letter *Y*, which gives a ghostly existence, and thus necessitates the performance of many expensive rites to counteract so undesirable a fate.' Possibly a similar chicanery was practised in Roman times! It would seem a reasonable inference anyway that the mystery and awe which the religious ceremony would lend to the casting of lots for purposes of divination would prevent the thinking person from speculating too deeply about it. Any attempt to try to forecast the result of a throw could undoubtedly be interpreted as an attempt to forecast the action of the deity concerned, and such an act of impiety might be expected to bring ill luck in its train. In addition, as we have noted, a method for such forecasting could not easily be made owing to the imperfections of most of the dice. On the other hand, it is possible that probabilities were known to the priests since the ceremonial dice are well made.

13. Through the Dark Ages the Christian church appears to have carried on guerilla warfare against gaming with knucklebones and dice. The writers of the Renaissance make many references to bishops who write *de aleatoribus* or *contra aleae ludum* during the first fourteen hundred years of the Christian era. It is likely therefore that the bishops wished to get rid of the sortilege as a religious ceremony, and they succeeded to a certain extent in doing this, although divination by lot still survives to-day in the Moravian sect. What the bishops could not do was to stop men playing games of chance. There are several references in early French literature to gaming. The play of Jean Bodel, *Le Jeu de Saint Nicolas*, written *c.* A.D. 1200, has a scene where thieves are gambling in a tavern. They are playing

* L. A. Waddell, *The Buddhism of Tibet*, W. Heffer and Sons Ltd. 1934 (2nd edition).

the dice game of 'Le hasard',* the rules of which were set down clearly by Pierre-Raymond Montmort in his book some five hundred years later. (*Analys sur le jeux d'azard* 1708, p. 113). Bodel's play is interesting for the suggestion that the thieves knew how to manipulate the dice to produce a desired result.

14. With the invention of printing (c. 1450) and its rapid development during the latter half of the fifteenth century the references to games of chance become more numerous, but there seems to be no suggestion of the calculation of probabilities. Thus we find in the writing of François Rabelais—a man who might be expected to know the latest in games of chance as played in taverns—the following interesting passage: 'Then they studied the Art of painting or carving; or brought into use the antic play of tables, as Leonicus hath written of it or as our good friend Lascaris playeth at it'† (*Gargantua and Pantagruel*, Urquhart's translation, Book I, Chapter XXIV). *Gargantua and Pantagruel* was issued in sections at intervals between 1532 and 1552. The date of this reference will therefore be not long after 1532.

The Leonicus of the reference is Nicolas Leonicus Tomeus, a professor of Greek and Latin at Padua who was born at Venice in 1456. He was well known for his learning and philosophical bent and acted as tutor to the English Cardinal Pole when as a young man he visited Italy. According to Erasmus he was 'a man equally respectable for the purity of his morals and the profundity of his erudition'. His letters, which have been translated by Cardinal Gasquet, give an interesting picture of the life of an intellectual of that time. He died at Padua in 1531 and his collected works were printed at Basel in 1532. Rabelais is clearly referring to Leonicus's treatise *Sannutus, sive de ludo talario*, a dialogue in the manner of Plato concerning the game of knucklebones (astragali). There is, however, little relevant to the calculus of probability in this work. The discussion turns on references to the game in Roman literature and a description and argument of the value of the various types of throw.‡

A similar type of disquisition was written by Calcagnini about this time but possibly a little later than that of Leonicus. Celio Calcagnini was born at Ferrara in 1479 and died there in 1541. He was a poet, a philosopher and astronomer of repute; his treatise *Quomodo caelum stet, terra moveatur, vel de perenni motu terrae commentatio*, in which he held that the earth moved round the sun, anticipated Galileo Galilei by some years, for Galileo was not born until 1564. The dissertation of Calcagnini entitled *De talorum, tesserarum ac calculorum ludis ex more veterum* is less philosophical in tone than that of Leonicus. It is of interest to probabilists only in that it was an influence over Cardano, who, from his several references, had clearly studied it closely.

* According to the editor, F. J. Warne, of the text of the play, *Le Jeu de Saint Nicolas*, 'hasart' meant the throw of a certain number of points at dice, varying according to the game played. In present-day probability theory the meaning is of course much wider.

† Rabelais actually wrote 'en usage l'anticque jeu des tables ainsi qu'en a escript Leonicus'. Duchat in the commentary on the 1741 edition says 'Ce n'est point *tables* qu'il faut lire ici, comme dans toutes les Editions, mais *tales*'. Presumably Duchat (followed by later commentators) makes this correction because of the work of Leonicus referred to. It is just possible that Rabelais meant what he wrote and that he was referring to the ancient board game (from which the modern game of backgammon developed) in which the 'men' may have been moved by throwing astragali, the counting of the throws being that described by Leonicus.

‡ I have not been able to trace why Lascaris is mentioned by Rabelais. Andre-Jean Lascaris surnamed Rhyndaconus (1445–1535), a Greek scholar born in Phrygia, was Librarian to François I. He rescued many Greek manuscripts from the Turks. Possibly he collected references to gaming in Greek literature much as Leonicus did for Roman?

15. We arrive at the sixteenth century, then, with a well-known humanist Leonicus, and a great astronomer Calcagnini, writing on games of chance with no attempt or reference to the calculation of a probability. (This does not mean of course that some calculations had not been made in a manuscript which we do not know about.) There were, moreover, other scholars and bishops writing on the same topic about this time, so that interest in the subject was keen. As far as we know at present it was left to Gerolamo Cardano to make the step forward. Cardano, the illegitimate son of a geometer, Fazio Cardano, was born in Pavia in 1501. His illegitimacy was a bar to his professional advancement on more than one occasion, and it is possible that the bitterness engendered by this fact was responsible for his not too scrupulous regard for the attribution of other scientists' ideas. The crime of plagiarism was a common accusation among scientific workers of the sixteenth and seventeenth centuries, but of none was it raised more loudly than of Cardano who was strongly disliked by his contemporaries and despised by his successors. Until about the middle of the nineteenth century his biographers unite in regarding him as a charlatan; possibly at the present time the pendulum has swung too far the other way, and more is read into his writings than is justified. The truth would seem to lie somewhere between the extremes of charlatan and persecuted savant.

Cardano was physician, philosopher, engineer, pure and applied mathematician, astrologer, eccentric, liar and gambler, but above all a gambler. He himself owns that on one occasion he sold his furniture and his wife's possessions in order to get money to indulge his passion for gaming, and there is no doubt that this passion was one of the things which ruled him through his whole life. His chief interest professionally was medicine, but he interested himself also in the communication of spirits and the casting of horoscopes. He does not seem to have been too successful at this last, but he was not deterred from casting that of Jesus, a performance the impiety of which probably led to his imprisonment. Even allowing for the exaggerations of his biographers there seems to be no doubt that he was eccentric to the point of madness. This did not prevent him, however, from making contributions to pure mathematics, and it is to this combination of pure mathematician and gambler that we owe the *Liber de Ludo Aleae*. This treatise was found in manuscript in Cardano's papers after his death in Rome in 1576, and was first published in his collected works in 1663 at Lyon. Cardano implies that it was written *c.* 1526; the exact date is not important since no question of priority or plagiarism is involved, but it is curious that a manuscript of this kind should have survived fifty years of his remarkably variegated career.

16. The first complete translation of *de Ludo Aleae* into English is given in *Cardano, the Gambling Scholar*, by Oystein Ore, published in 1953. Ore remarks that the book is badly composed and that understanding of Cardano's work has possibly been hindered by this. There are some, however, who will not agree with his commentary on the treatise and who may feel that as much prescience is now attributed to Cardano as there was before too little. The crux of Cardano's work is to be found in the section entitled 'On the cast of one die' in Ore's translation:

> The talus has four faces and thus also four points. But the die has six; in six casts each point should turn up once; but since some will be repeated, it follows that others will not turn up. The talus is represented as having flat surfaces, on each one of which it lies on its back; . . . and it does not have the form of a die. One half of the total number of faces always represents equality; thus the chances are equal that a given point will turn up in three throws, for the total circuit is completed

> in six, or again that one of three given points will turn up in one throw. For example, I can as easily throw one, three or five as two, four or six. The wagers are therefore laid in accordance with this equality if the die is honest....

We have therefore the necessary abstraction made; if the die is honest, i.e. if we may give equal weight to each side, then we may calculate the chances. There is no doubt, I think, that Cardano was led to this conclusion empirically and his generalization of it is partially wrong. For he goes on to discuss casts of two dice and three dice giving tables which are correct if 'the dice be honest'. When we come, however, to the section 'On play with knucklebones', it seems that he falls into error. The knucklebones (or astragali) have four sides. The different combinations of numbers which may arise in the throwing of four astragali are correctly enumerated, but the chances are calculated under the assumption that all sides are equi-probable; which they are not. Possibly Cardano had never played with astragali, for it is likely that if he had he would have noticed that to assume the sides of the astragalus had equal weight in his enumeration of alternatives was not adequate. But this fumbling suggests that he was not quite clear in his own mind about what he was proposing.

I do not think that the fact that Cardano did not quite see the mathematical abstraction clearly can detract from the fact that he did, on paper at any rate, as far as we know, calculate the first probability by theoretical argument, and in so doing he is the real begettor of modern probability theory. The claims of his biographer that he anticipated the law of large numbers, etc., may not be acceptable; it would appear that Cardano was judging from his experience rather than his algebra.

17. It would be strange if Cardano, following the mode of his age, did not communicate some of his thoughts about gaming to his pupils. Fear of being accused of plagiarism, fear of being plagiarized, may have kept him silent, but the whole tone of his treatise is a practical one; practical advice about playing, laying odds and so on make up a large portion of it. He would therefore almost certainly have discussed its contents with his friends, particularly if he thought about it over as long a period of time as he suggests. The fact that *de Ludo Aleae* did not appear in print until 1663 does not therefore seem to be a reason why Cardano's ideas should not have been common knowledge to scholars in Italy after his death, and the way in which Galileo-Galilei plunges into *his* discussion of dice playing, without much preamble, tends to lend colour to this.

Galileo-Galilei was born in Pisa in 1564, the son of Vincent Galilei, a musicographer well known in his day. He died in 1642 at Arcetri after a career as full of achievement as any that has ever been known. His contributions to science, both as astronomer and as mathematician, are striking for their originality of thought and clarity of purpose. Why this prince of scholars has never received the full recognition which is his due it is difficult to say. It is thought by some modern writers that his sensible recantation of the earth's movement, after physical torture at the hands of the Inquisition at the age of 70, has caused a revulsion to him among the scientists of later years. This is probably not so; what is more likely is that the envious fellow-scholars who delivered him to the Inquisition conspired after his death to belittle the work which he had done. In this they were possibly helped by Galileo's literary style which is noteworthy for clarity but not brevity, being in fact prolix and tedious in the extreme; no i is left undotted, no t is left uncrossed.*

* E. S. Pearson suggests to me that this prolixity was one which Galileo shared with many other Renaissance writers, and that it arose from the struggle which the early mathematicians must have had to formulate mathematical abstractions on paper. I think that this may well be so.

This being so if there was any doubt about the general method of procedure in calculating chances with a die we should have had a long disquisition on the subject. However, in *Sopra le Scoperte de i Dadi** he plunges straight away into his argument. The problem† is one already touched on by Cardano. Three dice are thrown. Although there are the same number of three partitions of 9 as there are of 10, yet the probability of achieving 9 in practice is less than that of throwing 10. Why is this? I quote a little from E. H. Thorne's translation of this note. The note begins:

> The fact that in a dice game certain numbers are more advantageous than others has a very obvious reason, i.e. that some are more easily and more frequently made than others, which depends on their being able to be made up with more variety of numbers. Thus a 3 and an 18, which are throws which can only be made in one way with 3 numbers (that is, the latter with 6, 6, 6 and the former with 1, 1, 1, and in no other way) are more difficult to make than, for example, 6 or 7, which can be made up in several ways, that is a 6 with 1, 2, 3 and with 2, 2, 2 and with 1, 1, 4 and a 7 with 1, 1, 5; 1, 2, 4; 1, 3, 3; 2, 2, 3. Again, although 9 and 12 can be made up in as many ways as 10 and 11 and therefore they are usually considered as being of equal utility to these, nevertheless it is known that long observation has made dice players consider 10 and 11 to be more advantageous than 9 and 12.

This extract serves to show how he begins the topic assuming that the calculations are known; it also serves to illustrate the prolixity of Galileo's style. After some discussion of the six 3 partitions of 9 and of 10, he goes on:

> Since a die has six faces and when thrown it can equally well fall on any one of these, only six throws can be made with it, each different from all the others. But if together with the first die we threw a second, which has also six faces, we can make 36 throws each different from all the others, since each face of the first die can be combined with each of the second....

After saying that the total number of possible throws with three dice are 216, he gives a table of the number of possible throws for a total of 10, 9, 8, 7, 6, 5, 4, 3, noting that the numbers 11–18 inclusive are symmetrical with these. Thus the number of possible throws for 10 is 27 and 25 for 9. His treatment of the problem is exactly that which we should use to-day and leaves us in no doubt that the calculation of a probability from the mathematical concept of the equi-probable sides of the die was clearly known to the sixteenth-century mathematicians of Italy. We can marvel at the person asking Galileo the question; he obviously gambled sufficiently to be able to detect a difference in empirical probabilities of 1/108.‡

18. Galileo's collected works were first published in Bologna in 1656, but this fragment on gambling was not included. It does appear in the more complete collection published at Florence in 1718. Since, however, Galileo thought the problem of little interest, for he did not pursue it, there seems to be no reason why he should have made a secret of it, and following the custom of his day he probably instructed his pupils. At any rate it is evident that the mathematical probability set was no stranger to the French mathematicians of the seventeenth century, as is witnessed by the now famous correspondence between Pascal and Fermat in 1654. The first letter of the series, from Pascal to Fermat, setting out the problem of points is missing. We have, however, Fermat's reply to it, and the subsequent

* This is Galileo's own title. *Considerazione sopra il Giuco dei Dadi*, a later title, appears first in the collected works of 1718.

† Like Pascal sometime later, Galileo wrote to answer a problem put to him by a gambler.

‡ M. G. Kendall points out to me that the problem posed by the Chevalier de Méré to Pascal concerning the problem of points involved similar small probabilities.

follow-up,* and from the way in which Fermat writes it seems clear that the actual definition of probability is assumed known. What the two savants were interested in was the application of this definition to specific problems which were concerned with dice playing between gamblers of equal skill and opportunity. The approach to the problems is similar to that of Galileo, and the generalizations which are made from the particular cases discussed are not well supported.

It is true that Galileo wrote on one problem only and fairly briefly at that, but it is difficult to see why Pascal and Fermat should be preferred as the originators of probability theory before Galileo or Cardano. It may well be that the precocity of Pascal as a mathematician led to much of his work being accepted with acclamation, and certainly without its priority being questioned. We find, for example, the famous Arithmetic Triangle in Stifel's *Arithmetic* (1543), in the *General Tratato* of Tartaglia in 1556, in the *Arithmetic of Simon Steven of Bruges* (Leiden, 1625). It is possible that Pascal may not have known of these writers. However, he certainly knew of Pierre Herigone's *Cours Mathematique* (Paris, 1634), since he makes several references to it in his own *Usage du Triangle Arithmétique pour trouver les puissances des binômes et Apotômes*. Herigone uses a table of numbers analogous to the Arithmetic Triangle to find binomal coefficients. Perhaps this same aura which dazzled Pascal's contemporaries (and at the same time caused them to overlook some of Fermat's work) still blinds us to-day.

19. If we take the origins for granted and look at developments of the theory, then by far the greatest impetus to theory during the years 1650–60 must have come from the publication of *De Ratiociniis in Aleae Ludo* by Christian Huygens. Huygens as a young man of 26 arrived in Paris in July 1655 on the equivalent of the English 'Grand Tour'. He did not meet Pascal, Fermat or Carcavi, the intimate friend of Pascal, but he did meet Roberval, professor of mathematics at the Collège Royal de France, who is mentioned by Pascal as having been also approached by the Chevalier de Méré. Huygens stayed in Paris from July to November, and after his return to Holland he began a correspondence with both Carcavi and Fermat which lasted over a period of years. The young man's imagination was obviously fired by the discussions he had in Paris, and his mathematical ambitions stimulated by the immense activity of the group which some ten years later (1665) was to found the *Académie des Sciences*. He set himself to work, and in March 1656 he wrote to Prof. van Schooten that he had prepared a manuscript about dice games. Francis Schooten was professor of mathematics at Leyden and had been Huygens's teacher. He took the young Huygens's manuscript (which was written in his native language), translated it into Latin and published it as an appendix to his *Exercitationes Mathematicae* in 1657. (A French translation of this appendix can be found in *Oeuvres de Huygens*, tome 14, on 'Calcul des Probabilités' published by *La Société Hollandaise des Sciences* in 1920.) In this *Tractatus de Ratiociniis in Aleae Ludo* Huygens sets out in a systematic manner what he must have learnt in Paris and adds some results which he may have achieved himself.

In the letter to Francis Schooten he writes

> ...quelques-uns des plus Célèbres Mathématiciens de toute la France se sont occupés de ce genre de Calcul, afin que personne ne m'attribue l'honneur de la première Invention qui ne m'appartient

* It is interesting to see Pascal fall into the same kind of trap which caused D'Alembert such controversy. In discussing the game of heads and tails and the tossing of a coin D'Alembert argued that the probability of throwing a head with two tosses of a coin was 2/3. For we may have *TT*, *TH* or *H*—when we stop, the second throw being immaterial, since we have achieved what we want.

pas. Mais ces savants. . . ont cependant cachés leurs méthodes. J'ai donc dû examiner et approfondir moi-même toute cette matière à commencer par les éléments, et il m'est impossible pour la raison que je viens de mentionner d'affirmer que nous sommes partis d'un même premier principe. . . .

Accordingly Huygens begins by proving his basic propositions, deals at some length with the problem of points and then passes on to dice playing. His last proposition (XIV) has a familiar ring:

> If another gambler and I throw 2 dice turn and turn about with the condition that I will have won when I throw 7 points and he will have won when he throws 6, if I allow him to throw first, find my chance and his of winning.

His delineation of his fourteen propositions is admirably clear and concise, and it is no marvel that the tract was used by mathematicians as a reference book up to the time of James Bernoulli (who reprinted it) and beyond. Possibly by this crystallization of the ideas of the French mathematicians Huygens has earned the right to be regarded as the father of the probability theory.

20. After Huygens the interest of probabilists was not solely in gaming, although this interest did not die away for another hundred years or so. But with Huygens the new calculus seems fairly launched, and this is therefore a suitable point to make a break. There are many questions which one leaves unanswered. The drawings and paintings by palaeolithic man of himself are very rare, and there is probably no hope of finding pictures of his recreations. If he prized the astragalus as a toy it seemed a possibility that he might have carved or decorated it in some way, but I have not been able to find any record of this. But while we cannot pull aside the curtain from four hundred centuries the possibility does exist that the pre-historians may be able, one day, to take us back a little farther than the third millennium. The farther back one goes the more fragmentary the evidence, but the earliest dice found are described as being of 'well-fired buff pottery', and they certainly would not have been the first made.

The tantalizing period to the present writer is the period from the invention of printing to A.D. 1600. In this period we have two mathematicians only calculating probabilities, and yet this was in the immense intellectual ferment of the Italian Renaissance. It seems hardly possible that there were not other natural philosophers who attempted similar calculations, but such documents, if they exist, will only now come to light by chance.

The correspondence between the French mathematicians of the first half of the seventeenth century is almost complete, and presumably the possibility does exist here of finding further letters. They all seemed at one time or another to send letters to one friend under cover of letters to another, and such letters may conceivably still be ascribed to the wrong person. However, enough information does exist regarding the seventeenth-century mathematicians to make a coherent study, and if I appear to have done them scant justice it is because I find the period so interesting that I hope to write about it more fully on another occasion elsewhere.

Collecting information about dicing and gaming has been a hobby of mine for some time, and the list of persons who have drawn my attention to one aspect or another of it is formidable. I want to thank Prof. B. Ashmole of the British Museum who allowed me critically to examine the dice of the classical period which are in his care and M. Jean Charbonneaux of the Musée de Louvre who did me the same service. To Prof. C. M. Robert-

son of my own college I owe not only the privilege of tossing many dice but many stimulating discussions and useful references. A. J. Arkell allowed me to examine the dice brought by Prof. Sir Flinders Petrie from Egypt and to photograph various gaming boards not reproduced here. The breadth of knowledge and wide reading of Miss M. S. Drower have acquainted me with many Egyptian board games which provide a fascinating puzzle for those interested in deducing how they are played. Miss J. Lowe and R. Graves drew my attention to various references in classical literature. The illustration of the Hounds and Jackals game is by the courtesy of the Metropolitan Museum of Art of New York.

I want to thank Dr J. H. Willis who translated the *Sive de ludo talario* of Nicolas Leonicus for me, Dr E. H. Thorne who supplied a translation of Galileo's letter on dice, Prof. B. Woledge who drew my attention to early French plays and Miss J. Townend who drew Text-fig. 1. Miss J. Pearson and Miss J. Edmiston helped me to find many references and A. Munday and Miss A. Lodge helped with photographs. The manuscript as a whole owes much to the keen critical faculties of Prof. E. S. Pearson and Prof. M. G. Kendall. Part of this work was carried out with the aid of a grant from the Central Research Fund of the University of London.

SOME FEATURES OF THE GENERATION TIMES OF INDIVIDUAL BACTERIA

By E. O. POWELL

Microbiological Research Department, Experimental Station, Porton

Introduction

The dynamics of bacterial growth in the mass has been the subject of a great deal of experimental study, but quantitative observations on the behaviour of individuals are scanty. In spite of the primary need for detailed knowledge of normal growth processes, it is the abnormal and morbid that have received most attention from both experimenters and theoreticians. Kendall (1952*a*) and Finney & Martin (1951) have recognized the necessity for extended studies of the normal condition; in this they have been prompted by the existence of two hypotheses (Kendall, 1948, 1952*a*; Rahn, 1932), both of which connect the observed scatter of generation times with an inner mechanism of fission.

Measurements of individual generation times sufficiently systematic and detailed to make analysis worth while have been provided only by Kelly & Rahn (1932) so far as I know. My object here is, first, to add to the existing corpus of data; secondly, to show that for technical reasons, the experimental work so far done (including my own) is not really adequate to test the hypotheses that have been proposed; thirdly, to suggest further lines of study.

According to Rahn's hypothesis, the fission of a bacterium involves the simultaneous duplication of a number of essential entities or structures which may be the genes. The time required for duplication will not be the same for each, but will be subject to a law of 'chance', and fission can only take place when every entity has been duplicated. Under certain assumptions, Rahn shows that the observed generation times should then be scattered according to the law

$$dF = \frac{g}{m} e^{-\tau/m} (1 - e^{-\tau/m})^{g-1} d\tau, \tag{1}$$

where dF is the proportion of generation times in the range τ to $\tau + d\tau$. The parameter g is equal to the number of entities which have to be duplicated, and m is a parameter depending on the rate of the individual processes. In what follows, I call the frequency function (1) 'Yule's distribution' (since Yule (1925), was the first to describe it).

Kendall's reasoning is similar, except that he supposes the events leading to fission to take place step by step, and fission to occur only when the series of steps is complete. Each step may or may not be a process of duplication. The observed generation times should then be scattered according to a Pearson Type III law in the form

$$dF = \frac{\tau^{g-1} e^{-\tau/m}}{m^g \Gamma(g)} d\tau,$$

where g is the number of steps and m is a kinetic parameter as before.

The idea of a distribution of generation times implies in practice (though this is not strictly necessary) some degree of permanence in the properties of the population to which it

applies. At a given instant, the organisms in a bacterial culture certainly possess a definite age distribution, but this has no significance if the growth rate and cultural conditions have varied during the life span of the oldest extant organism. *A fortiori*, a distribution of generation times, even if it be formally defined at an instant, will be a complex and imperspicuous character of its population unless the factors controlling it remain constant over at least several generations. Conditions adequate to the precise definition and determination of a distribution probably exist only, if at all, in a continuous culture maintained in a steady state. No measurements have yet been made on such a system, and I have assumed here (as has been done implicitly in the past) that during the phase of logarithmic growth in static culture the succession of generations proceeds in a regular manner. I give, however, two examples to the contrary.

It is usually found that the mass growth rate (the logarithmic derivate of the mass of organisms per unit volume) in a static culture rapidly accelerates at the end of the lag phase to a steady value which is maintained for several hours. The number growth rate (the logarithmic derivate of the number of organisms per unit volume) often behaves less regularly. It may be small or zero even after the mass growth rate has reached its maximum value; abnormally long organisms are thereby formed. There follows a stage of acceleration during which the number of growth rate approaches and temporarily exceeds the mass growth rate. The two then become equal, in some cases only for a short time before the phase of decline sets in. These remarks apply to culture in a liquid medium. Detailed quantitative information about the growth of colonies on solid media over long periods (i.e. of many generations) appears to be lacking, and it is only on solid media that observations of individual generation times can be made. With these difficulties in mind, it is clear that a satisfactory experimental procedure is not easily attained. Observations which are to be combined so as to furnish a distribution curve must either be confined to a brief period in the life of each culture examined, or, if extended, they must be shown to be drawn from a sensibly unchanging population.

In a culture whose number growth rate is varying there is temporarily an excess or deficit of young organisms; a change in the growth rate implies a change in generation time distribution. It does not necessarily imply a change in *all* the parameters of that distribution, but even if some one were known or assumed to be constant, it would be a matter of considerable practical difficulty to extract its value from a set of data derived from a changing population. At present it is not possible to test or to make use of Kendall's (1948) relation between the coefficients of variation of clone size and generation time; a method of maintaining a constant growth rate for many generations must first be established.

The results which I have to present consist of measurements of generation times carried out on six species of organisms (*Bacterium aerogenes*, strain N.C.T.C. 8197; *Bact. coli anaerogenes*, strain N.C.T.C. 4450; *Streptococcus faecalis*; *Proteus vulgaris*, strain LC; *Bacillus mycoides*, strain SR 2; *B. subtilis*, strain UP 1). Their distributions are compared with the frequency functions suggested by Kendall and Rahn. A partial analysis and discussion of some other features of the nexus of generation times are also given. I use the words 'mother', 'daughter', 'sister' with an obvious extension of their usual meaning, and the neutral terms 'inception', 'termination' to denote respectively the events at which an organism becomes a recognizably separate entity (by fission of its mother) and by which it ceases to be so (by itself dividing). The same terms may also refer to the epochs of these events. In describing the morphological changes in the cell wall occurring at fission I replace

much more accurately since inception and termination took place in close juxtaposition both temporally and spatially; records of generation times of 2, 4 and 6 min. may be taken as correct to ± 1 min., and therefore also correct as to their frequency. As it turns out, this is an important consideration.

The uncertainty introduces an error of a special kind. A generation time which should be recorded as τ may be recorded as $\tau - 2$ or $\tau + 2$ for instance; it will then appear in a frequency group adjacent to the correct one, and so there will be correlated errors in the frequencies. The effect will increase the variance, probably to a quite negligible extent, but there may also be a concomitant error in the χ^2 between the observations and a fitted function, making the fit seem better or worse than it is. I have found no way of estimating or allowing for this aberration.

I wish to make it quite clear what is here meant by a measured generation time τ: The fission of an organism is first observed to be complete, as nearly as can be judged, at a time t_1, say. It has therefore divided at some time between $t_1 - 2$ and t_1. Similarly, one of its daughters terminates at some time between $t_2 - 2$ and t_2. Then the generation time of that daughter is recorded as $t_2 - t_1$, which is always a multiple of 2. Thus τ is the mean value of the generation time of the group of organisms to which it is ascribed. I make this simple point because a careful study of Kelly & Rahn's (1932), Rahn's (1932) and Finney & Martin's (1951) papers raises doubts as to what Kelly & Rahn have actually recorded. Their figs. 1 and 2 suggest that they proceeded as I have done (though with a 5 min. instead of a 2 min. interval between observations), but in their tables the observations are grouped into ranges 0–5, 5–10, 10–15, etc. Rahn and Finney & Martin have the misleading phrase 'fissions ... observed in successive five-minute intervals'. Finney & Martin assumed that the mean generation times for these intervals were $2\frac{1}{2}$, $7\frac{1}{2}$, $12\frac{1}{2}$, ...; they may have been $2\frac{1}{2}$, 5, 10, 15, ... or 5, 10, 15, 20, ... I have verified if either of these alternatives is chosen, the goodness of fit of Yule's distribution is about the same as that obtained by Finney & Martin ($\chi^2_9 = 14{\cdot}9$), but the corresponding estimates of g are widely and significantly different: 18·2 and 36·2 respectively, as against 26·0.

Other observational arrangements are possible. For instance, the distribution of generation times can be deduced from the age distribution in a culture. If $\phi(a)$ is the frequency function of age and $f(\tau)$ that of generation time, it is found that if the culture is growing steadily

$$\phi(a) = \phi(0)\, e^{-ka} \int_a^\infty f(\tau)\, d\tau \qquad (4)$$

(cf. Harris, 1951). Unfortunately, the form of the curve $\phi(a)$ is dominated by the value of k; it is very insensitive to changes in the parameters of $f(\tau)$ that leave k the same. Experimentally, the determination of an age distribution at a given time requires a study of the previous history of the culture over a period long enough to give the generation time distribution directly. I have, however, made use of the relation (4) in the reverse sense, as a check on the regularity of growth in experiments with *B. mycoides* and *B. subtilis*.

Again, suppose that a number of organisms extant at the epoch $t = 0$ are watched until they divide. The times $t = \theta$ of fission will be distributed with a frequency function

$$\phi(0)\, e^{k\theta} \int_\theta^\infty e^{-k\tau} f(\tau)\, d\tau\, d\theta.$$

This gives another possible method of determining $f(\tau)$; I have made no use of it.

(*b*) *Practical details*

With one exception, organisms were grown on cellophane over tryptic meat broth in the culture chamber described by Harris & Powell (1951). The medium was constantly circulated so that an approximation to a condition of continuous culture obtained; the volume of medium (100 c.c.) was so large relative to the inoculum used (*c.* 100 organisms) that no appreciable change could occur during the course of an experiment, either by depletion of nutrients or accumulation of products of metabolism. In this respect the technique is superior to that of Orskov (1922) adopted by Kelly & Rahn. On the other hand, times of fission can be judged with less precision in cellophane culture; the average error appeared to be about $\pm$ 2 min. However, this contribution to the total variance is only a small fraction of that due to the organisms themselves.

The cultures were incubated throughout at a temperature of $35 \pm 0{\cdot}5°$ C. The two *Bacillus* species were inoculated as spores, the rest as saline dilutions from a 20 hr. liquid culture. To avoid confusion, the organisms on which observations were to be made were separated from their neighbours by means of a micromanipulator so that each lay within a clear area of cellophane. They were exposed to light as little as possible. Evaporation from the surface of the cellophane was controlled so that each organism or group of organisms was surrounded by a visible fillet of liquid (Harris & Powell, 1951). The evaporation was temporarily increased only when necessary to verify the occurrence of fission. A photographic method of recording was rejected after trial as less informative than direct examination, and no less tedious when the processing and assessment of films were taken into account.

No difficulty was encountered in working with *Bact. aerogenes* and *Bact. coli anaerogenes*. A typical experiment consisted in following the development of four organisms simultaneously for two or three complete generations. Organisms were selected for observation always at about 2 hr. from inoculation, previous experience having shown that erratic behaviour at the end of the lag phase would by then have subsided.

Strep. faecalis was found to be very sensitive to light; a few minutes exposure to the usual conditions of illumination sufficed to stop growth. The effect was minimized by adding 5% of sheep serum to the medium and illuminating with red light as dim as could be tolerated. As a further precaution, observation was confined to one complete generation only of the progeny of the selected organisms. The numerical results suggest that these measures were effective in securing uniform unrestricted growth over the short period necessary. Observation was begun at $2\frac{1}{2}$ hr. from inoculation.

The mode of fission of *Pr. vulgaris* in cellophane culture was erratic, and seemed sometimes to be intermediate in character between the septate and the isthmoid. The first signs of a waist were apparently accompanied by the formation of a septum, and the organisms would remain for several minutes in this state of incipient fission. Attempts to assess generation times were unsatisfactory, and this species was therefore studied in culture on an agar-tryptic meat medium under incident dark-field illumination (Pearce & Powell, 1951). Judging by the lengths of organisms as then seen, the uncompleted fissions observed by the vertical illumination on cellophane were recorded as complete fissions under the dark field. For this organism, then, the conditions of culture were similar to those in Kelly & Rahn's experiments. For $3\frac{1}{2}$–4 hr. from inoculation it appeared to grow and reproduce steadily, but then the rate of fission declined, longer motile organisms were formed and swarming began. In order not to encroach on this phase of changing growth rate, observations were

again confined to one generation only, beginning at 2½ hr. from inoculation. (This organism did not swarm on cellophane over a circulating medium, presumably because no local concentration of metabolic products could be built up. See Lominski & Lendrum, 1947.)

Observations on *B. mycoides* were carried out during the 80 min. beginning at 2½ hr. from inoculation. By that time the spore coat had been shed and families of two to four organisms developed. Three groups of experiments were conducted with *B. subtilis*: series (i), observation begun at 2½ hr.; series (ii), observation begun at 4½ hr.; series (iii), organisms growing over defibrinated sheep blood diluted with its own volume of normal saline; observation begun at 3 hr. In order to check the regularity of growth of these two species, the ages of the organisms extant at the end of each experiment were recorded, so that their distributions could be compared with those calculated by means of (4) from the generation time distributions.

The frequency functions

I give in this section a statement of the frequency functions which are made use of below, together with their relevant properties. Each is limited to positive values of the variate, and each involves a parameter g which, whatever its interpretation, is a measure of intrinsic dispersion, i.e. a function of the coefficient of variation only.

Fitting was carried out either by the method of maximum likelihood or by that of moments (sometimes both). In several of the examples the distributions are such that the method of moments is of low efficiency (cf. Fisher, 1922). However, I do not think that the results are of much more than comparative value in any case, and the points I shall have to make are not so nice as to call for great statistical refinement.

(*a*) *The Pearson Type III distribution* (*Kendall's hypothesis*)

$$dF = \frac{\tau^{g-1} e^{-\tau/m}}{m^g \Gamma(g)} d\tau.$$

This is a well-known function. The cumulants are

$$\kappa_r = m^r g \Gamma(r),$$

whence the first three moments

$$\mu_1' = mg, \quad \mu_2 = m^2 g, \quad \mu_3 = 2m^3 g.$$

The coefficient of variation is

$$c_0 = g^{-\frac{1}{2}}$$

and the skewness and kurtosis

$$\gamma_1 = \kappa_3/\kappa_2^{\frac{3}{2}} = 2g^{-\frac{1}{2}}, \quad \gamma_2 = \kappa_4/\kappa_2^2 = 6/g.$$

Both tend to zero for large g.

Maximum-likelihood estimates $\hat{g}$, $\hat{m}$ of g and m are given by

$$\Sigma f\tau/n - \hat{g}\hat{m} = 0,$$

$$\Sigma f \log \tau/n - \log \hat{m} - \psi(\hat{g}) = 0,$$

where the f are observed frequencies of the generation times τ, n is Σf and $\psi(\hat{g})$ is

$$d \log \Gamma(\hat{g})/d\hat{g}.$$

(b) Yule's distribution (Rahn's hypothesis)

$$dF = \frac{g}{m} e^{-\tau/m}(1-e^{-\tau/m})^{g-1}\,d\tau.$$

Except for μ_1', the moments are difficult to calculate and are best obtained through the cumulants (cf. Daniels, 1952; but see Yule, 1925). The cumulant generating function is easily found to be

$$K(t) = \log\Gamma(g+1)+\log\Gamma(1-itm)-\log\Gamma(g+1-itm),$$

whence
$$\kappa_r = (-1)^r m^r\{\Gamma_r(1)-\Gamma_r(g+1)\},$$

where for uniformity I write

$$\Gamma_r(x) = (d/dx)^r\{\log\Gamma(x)\}, \quad \Gamma_1(x) \equiv \psi(x).$$

In particular,
$$\mu_1' = \kappa_1 = m\{\Gamma_1(g+1)-\Gamma_1(1)\}, \tag{5}$$

$$\mu_2 = \kappa_2 = m^2\{\Gamma_2(1)-\Gamma_2(g+1)\}, \tag{6}$$

$$c_0 = \frac{\{\Gamma_2(1)-\Gamma_2(g+1)\}^{\frac{1}{2}}}{\Gamma_1(g+1)-\Gamma_1(1)},$$

$$\gamma_1 = \frac{\Gamma_3(g+1)-\Gamma_3(1)}{\{\Gamma_2(1)-\Gamma_2(g+1)\}^{\frac{3}{2}}}.$$

The curve is distinguished from a Type III having the same first two moments by its slower rise near zero, and by its greater steepness on the left flank. In fact, for increasing g, γ_1 diminishes but tends to a still rather large non-zero limit

$$-\pi^{-3}6^{\frac{1}{2}}\Gamma_3(1) = 1{\cdot}14.$$

In practice γ_2 scarcely differs from $2{\cdot}4$ $(= 3!(\pi^4/90)(6/\pi^2)^2)$.

Some other features are discussed by Finney & Martin (1951) and Rahn (1932), who note the insensitivity of the form of the curve to changes in g.

The maximum-likelihood equations for fitting are

$$\frac{n}{\hat{g}} + \Sigma f \log(1-e^{-\tau/\hat{m}}) = 0,$$

$$-\frac{n}{\hat{m}} + \Sigma f\tau/\hat{m}^2 - \frac{\hat{g}-1}{\hat{m}}\Sigma\frac{f\tau/\hat{m}}{e^{\tau/\hat{m}}-1} = 0.$$

The solution is greatly facilitated by the tables of Einstein functions due to Sherman & Ewell (1942).

(c) The Pearson Type III distribution with allowance for bias

From equation (3), if the true distribution is of Pearson Type III, that derived from data which are biased in the way already described will be

$$dF = B\alpha(\tau)\frac{\tau^{g-1}e^{-\tau/m}}{m^g\Gamma(g)}d\tau, \tag{7}$$

where $\alpha(T)$ is given by equation (2). This is a frequency function in the range $0 \leqslant \tau \leqslant T$ ($T = 80$ in all experiments); the moments strictly involve incomplete gamma functions, but

these are here so nearly unity that no appreciable error arises in taking the range of integration as infinite. And then it is found that

$$B = \left\{\frac{e^{kT}(1+km)^{-g}-1}{e^{kT}-1}\right\}^{-1},$$

$$\mu_1' = mg\left\{\frac{e^{kT}(1+km)^{-g-1}-1}{e^{kT}(1+km)^{-g}-1}\right\},$$

$$\mu_2' = m^2g(g+1)\left\{\frac{e^{kT}(1+km)^{-g-2}-1}{e^{kT}(1+km)^{-g}-1}\right\},$$

with $k = (\log 2)/(mg)$. The parameters are hence obtained by iteration; approximate values of m and g are inserted into the expressions in curly brackets, and the two equations for the moments solved to give a closer approximation which is used as the starting point for a second cycle, and so on. For the *terminus a quo*, it is enough to take $m = 1$, and $\bar{\tau} = gm = 25$ min. (a value roughly known *a priori*).

(d) Yule's distribution with allowance for bias

$$dF = B\alpha(\tau)\frac{g}{m}e^{-\tau/m}(1-e^{-\tau/m})^{g-1}d\tau. \tag{8}$$

With the same approximation as in the previous paragraph,

$$B = (e^{kT}-1)/H,$$

$$\mu_1' = m\{\Gamma_1(g+1)-\Gamma_1(1)\}\left[H\frac{\Gamma_1(1+km+g)-\Gamma_1(1+km)}{\Gamma_1(g+1)-\Gamma_1(1)}-1\right]\Big/(H-1),$$

$$\mu_2' = m^2[\Gamma_2(1)-\Gamma_2(g+1)+\{\Gamma_1(g+1)-\Gamma_1(1)\}^2]$$
$$\times\left[H\frac{\Gamma_2(1+km)-\Gamma_2(1+km+g)+\{\Gamma_1(1+km)-\Gamma_1(1+km+g)\}^2}{\Gamma_2(1)-\Gamma_2(g+1)+\{\Gamma_1(g+1)-\Gamma_1(1)\}^2}-1\right]$$
$$\times[H-1]^{-1},$$

where
$$k = (\log 2)/m\{\Gamma_1(g+1)-\Gamma(1)\},$$
$$H = e^{kT}\frac{\Gamma(g+1)\,\Gamma(1+km)}{\Gamma(g+1+km)}.$$

After fitting the function (7) to a set of data, estimates of the true moments were calculated from its parameters, and were used to obtain values of the parameters of (8) by means of (5) and (6). These values were taken as the first approximation in fitting (8), iteration being carried out as in the previous section.

Analysis of the data

General

The fission of a micro-organism, as usually understood, consists in its separation into two geometrically distinct pieces, each with a closed cell wall separating it from its surroundings. The closure of the wall follows the division of the cytoplasm at some unknown interval, but in unicellular organisms there is at least a unique one-to-one correspondence between the true cell division and the observed fission. In the genus *Bacillus* this is not so; the organisms are multicellular, and at any time contain party walls of various degrees of completeness.

The party walls may or may not be laid down as double layers, but they cannot be seen to be so, and it is not known that each becomes necessarily a site of fission; in these organisms fission is a tertiary rather than a secondary process (for recent work on the structure of the party wall, see Dawson & Stern, 1954). Thus the generation times of *B. mycoides* and *B. subtilis* have not the significance for the hypotheses of Rahn and Kendall that those of the other organisms have. (The same applies to Kelly & Rahn's measurements on *B. cereus*; Rahn recognized this.) There is no *a priori* reason why either hypothesis should not apply, as regards its mathematical form, to the *Bacilli*, but the 'genetic' parameter g cannot then be interpreted as the number of genes or primitive steps determining fission. For the moment, little more than a descriptive treatment of the generation times of these organisms can be usefully attempted in this context.

The raw data are set out in Table 1, together with the first two moments and the skewness and kurtosis of each distribution. The range of generation times is very great, and much at variance with the vague assumption often made, that bacteria divide rather regularly. A striking feature in all the experiments was the highly erratic behaviour of small families of organisms (e.g. the six or fourteen constituting the progeny, up to two or three generations, of a single ancestor); in some the generation times are quite uniform, in others widely dispersed. This unforeseen lack of homogeneity makes statistical analysis difficult, and renders doubtful the significance of any broad and simple treatment. Such quantitative analysis as I have been able to apply can only be a makeshift; but a more refined investigation will not be justified until a larger range of data has accumulated.

In the paragraphs which follow, some forward references are unavoidable; homogeneity is considered first because it affects the methods adopted in assessing the parameter g, but the results of curve-fitting are also relevant, and have been used to modify the tests of homogeneity.

Homogeneity of the data

Because of the special difficulties associated with *B. mycoides* and *B. subtilis*, only the four unicellular organisms *Bact. aerogenes*, *Bact. coli anaerogenes*, *Strep. faecalis* and *Pr. vulgaris* will be considered in this section.

Kendall (1948) by applying Bartlett's (1937) test for homogeneity of variance, showed that the estimates of his g obtained from the several experiments of Kelly & Rahn were highly inconsistent; and later (1952*a*) he expressed doubts as to the propriety of combining the data to obtain an overall estimate of g. Such doubts are of course justified if the source of variation is a change in the experimental conditions, and certainly Kelly & Rahn's results strongly suggest that day-to-day fluctuations are considerable.

Any one experiment of the present series yields measurements of the generation times of organisms belonging to several families, each the progeny of a single organism selected at the beginning of the experiment. The organisms all lie within a single field of the microscope and are subjected therefore to nearly identical environmental conditions. The variance of generation times within the experiment must be that proper to the organism under those conditions, and the coefficient of variation (a constant for each species by hypothesis) should yield an unbiased measure of g.

Now it remains true of the present results, as well as of Kelly & Rahn's, that there is apparent lack of homogeneity between experiments. An examination of the coefficients of variation showed them to be fairly uniformly distributed, except for a few which were outstandingly large; the principal contribution to the variance in these cases was provided by

Table 1. *Frequency of generation times, with the first (μ_1', $=\bar{\tau}$) and second (μ_2) moments and the measures γ_1, γ_2 of skewness and kurtosis*

Generation times	*Bact. aerogenes*	*Bact. coli anaerogenes*	*Strep. faecalis*	*Pr. vulgaris*	*B. mycoides*	*B. subtilis* series (i)	*B. subtilis* series (ii)	*B. subtilis* series (iii)
0	—	—	—	—	—	—	—	—
2	—	—	—	—	2	4	1	1
4	—	—	—	—	4	6	2	2
6	2	3	1	4	6	8	2	3
8	3	8	0	4	10	15	6	2
10	4	7	0	2	12	13	3	3
12	13	24	1	5	19	26	13	6
14	35	45	4	9	13	24	14	15
16	48	49	6	7	20	32	23	27
18	75	56	13	20	25	34	24	35
20	72	47	8	34	22	42	25	25
22	65	53	15	42	27	27	25	40
24	40	39	9	42	23	39	34	44
26	42	30	20	41	17	19	30	32
28	29	17	12	29	12	22	32	28
30	9	16	11	34	9	13	28	28
32	13	8	8	26	14	15	19	18
34	2	3	4	25	5	17	24	14
36	4	4	3	17	6	8	6	12
38	1	7	0	19	11	1	11	12
40	1	0	0	13	9	1	11	11
42	2	4	1	1	4	3	7	9
44	0	1	0	4	3	—	3	1
46	2	—	0	3	4	—	1	5
48	—	—	1	1	1	—	1	1
50	—	—	0	1	1	—	2	3
52	—	—	1	0	0	—	2	0
54	—	—	—	2	0	—	1	1
56	—	—	—	2	1	—	2	1
58	—	—	—	3	0	—	1	0
60	—	—	—	—	0	—	—	2
62	—	—	—	—	0	—	—	—
64	—	—	—	—	1	—	—	—
66	—	—	—	—	0	—	—	—
68	—	—	—	—	0	—	—	—
70	—	—	—	—	0	—	—	—
72	—	—	—	—	0	—	—	—
74	—	—	—	—	1	—	—	—
76	—	—	—	—	—	—	—	—
78	—	—	—	—	—	—	—	—
80	—	—	—	—	—	—	—	—
Σf	462	421*	118	390	282	369	353	381
No. of expts.	14	12	9	19	16	19	18	16
No. of families measured	51	51	59	93	49	43	53	66
μ_1'	21·05	20·48	24·90	27·25	22·90	20·38	25·72	25·64
μ_2	32·94	43·98	46·37	75·43	119·76	66·40	89·64	83·99
γ_1	0·79	0·72	0·67	0·57	0·87	0·18	0·45	0·66
γ_2	1·78	0·84	2·07	1·25	1·60	−0·46	0·70	1·12

* One organism failed to grow and suffered lysis.

one or (rarely) more families of very unequal generation times. Thus for example with *Bact. coli anaerogenes* in an experiment involving four families each of six organisms, the mean generation times were

26·3, 21·7, 28·3, 27·0;

the variances 51·8, 2·2, 56·6, 135·5;

and the squared coefficients of variation

0·0748, 0·00467, 0·0707, 0·186.

Irregularities of this kind, which were found in all the species examined, appear to make it doubtful that any simple meaning can be attached to the coefficient of variation of larger samples.

In order to test the homogeneity of within-family variance, the Bartlett criterion, often denoted by M (which for normal samples is distributed approximately as a χ^2), was computed for each experiment in the usual way and the sum, ΣM, formed for each group of experiments. A correction for kurtosis was applied, given by Box (1953) as

$$\Sigma M' = \Sigma\left\{M \Big/ \left(1+\frac{1}{2}\frac{\nu}{\nu+1}\gamma_2\right)\right\}.$$

This expression was simplified by taking a common corrective factor $(1+\frac{1}{2}\gamma_2)$ for every family, so biasing M' towards smaller values.

Two assumptions were made in assigning γ_2: (i) If the underlying distribution is of Type III, γ_2 is $6/g$. In this case the γ_2 were calculated from the final estimates of Kendall's g given by curve fitting (Table 9). (ii) If the underlying distribution is Yule's, γ_2 is very nearly 2·4. The results of the test given in Table 2 show that the apparent heterogeneity is not significant under Rahn's hypothesis, but is markedly so under Kendall's. To this extent it might be argued that Rahn's hypothesis is to be preferred. But examination of the figures in detail shows that the large contributions to M' come mainly from families containing members of unusually short generation time, whereas it is the long upper tail of Yule's distribution which inflates the fourth moment; near the origin the ordinates are exceedingly small. There is, however, no sufficient reason here to suspect Kelly & Rahn's experiments of being ill-controlled.

In order to assess the extent to which the overall variance of generation time is increased by inconstant experimental conditions, I have carried out conventional analyses of variance on the means of families. From Table 3 it can be seen that in every case a significant part of the total variance is contributed by differences in the growth rate from experiment to experiment. If σ_F^2 and σ_E^2 are the true within and between experiment variances, the observed between-experiment mean square will be an estimate of

$$\sigma_F^2+\frac{N^2-\Sigma n^2}{N(K-1)}\sigma_E^2$$

(where K is the number of experiments, N the number of families, and n the number of families in an experiment). On this basis we derive the figures in the last column of Table 3 as estimates of the experimental variance of the growth rate. Except for *Pr. vulgaris*, these estimates of σ_E^2 are not great; nevertheless, estimates of the coefficient of variation from the raw data will be too high.

Table 2. *Modified Bartlett test of homogeneity of variance of families*

	Bact. aerogenes	*Bact. coli anaerogenes*	*Pr. vulgaris*
(A) Kendall hypothesis: γ_2	0·38	0·55	0·43
$\Sigma M'$	75	78	109
n	28	39	72
$P(\Sigma M')$	0·000	0·000	0·000
(B) Rahn hypothesis: γ_2	2·4	2·4	2·4
$\Sigma M'$	41	45	60
n	28	39	72
$P(\Sigma M')$	c. 0·05	$>0·1$	$>0·1$

Table 3. *Analysis of variance of family means*

Organism	Source of variation	Sum of squares	Degrees of freedom	Mean square	F-ratio (and $P(F)$)	Estimated σ_z^2
Bact. aerogenes	Between expts.	305·95	13	23·53	2·03 (0·05)	3·30
	Within expts.	429·32	37	11·60		
	Total	735·27	50			
Bact. coli anaerogenes	Between expts.	314·16	11	28·56	2·86 (<0·01)	4·38
	Within expts.	389·34	39	9·98		
	Total	703·50	50			
Strep. faecalis	Between expts.	610·18	8	76·27	2·37 (<0·05)	6·86
	Within expts.	1609·21	50	32·18		
	Total	2219·39	58			
Pr. vulgaris	Between expts.	2037·45	18	113·19	3·58 (<0·0001)	17·23
	Within expts.	2274·42	72	31·59		
	Total	4311·87	90			

By contrast, the coefficients of variation of generation times of families, with the exception of *Pr. vulgaris*, are not significantly perturbed. The coefficients themselves (and their squares) are very unsymmetrically distributed, but the distribution of their logarithms, as shown by a graphical test, is roughly normal, and analysis of variance carried out on the logarithms yields the results of Table 4. (For convenience the variable was taken as $100 \log_{10} c^2$, to the nearest integer; the constant factors of course cancel in the F-ratio.) *Strep. faecalis* again had to be omitted from this test, because many families (each consisting of a pair of sisters only) had zero variance. However, the test was also carried out on the squared coefficients of variation themselves; the sense of the results was exactly the same, and *Strep. faecalis* gave an F-ratio probability above 0·1.

Table 4. *Analysis of variance of family coefficients of variation*

Organism	Source of variance	Sum of squares	Degrees of freedom	Mean square	F-ratio (and $P(F)$)
Bact. aerogenes	Between expts.	20386	13	1568	0·84 (–)
	Within expts.	67246	36	1868	
	Total	87632	49		
Bact. coli anaerogenes	Between expts.	17668	11	1606	0·86 (–)
	Within expts.	72701	39	1864	
	Total	90369	50		
Pr. vulgaris	Between expts.	159425	18	8857	2·94 (0·001)
	Within expts.	216582	72	3008	
	Total	376007	90		

Thus the coefficient of variation is relatively stable as between experiments, but its large interfamily variance is difficult to account for on Kendall's hypothesis if, as is supposed, the generation time of an organism is independent of its immediate ancestry. It may well be a product of 'delayed fission'; that is, the lapse of an appreciable interval between the division of nucleus or cytoplasm, and the separation of cell wall, A single large delay in a family postpones the termination of one organism and the inception of two others, and so may have a disproportionate effect on the coefficient of variation. To admit this possibility is to entertain the view that the measurements may not be directly relevant to either hypothesis.

The experiments with *Pr. vulgaris* were evidently less successful than the rest as regards reproducibility of growth rate; this is perhaps to be associated with the organisms' potentially more complex behaviour on the static medium used.

Correlation between generation times

Kelly & Rahn (1932) suggested that some of the more extreme values occurring among their measurements were to be accounted for by 'delayed division'. Such delay would obviously tend to generate negative correlation between mother and daughter and positive correlation between sisters. Rahn (1932) showed that this tendency existed, but he did not consider the measurements as a whole. It is probable that delayed division in this sense is too rigid a conception, and that a delay of variable magnitude always occurs; this does not affect the conclusion.

The product-moment correlation coefficients between generation times of mothers and daughters (ρ_{MD}) and pairs of sisters (ρ_{SS}) in the present series of measurements are given in Table 5. Not all the mother-daughter pairs are independent, since in a family of three or four generations some organisms appear twice as mothers and once as daughters. Scatter diagrams of the mother-daughter correlations showed the entries to be fairly uniformly

Table 5. *Correlation between generation times of mothers and daughters (ρ_{MD}) and sisters (ρ_{SS}). Number of pairs of observations, p. Parentheses indicate that the coefficient is not significantly different from zero at the 5% level*

Organism	ρ_{MD}	p	ρ_{MD} corr. for bias	ρ_{SS}	p	ρ_{SS} corr. for bias
Bact. aerogenes	(0·069)	360	—	0·77	231	—
Bact. coli anaerogenes	(−0·051)	318	—	0·48	210	—
Strep. faecalis	—	—	—	0·68	59	—
Pr. vulgaris	—	—	—	0·74	195	—
B. mycoides	−0·18	192	−0·42	0·42	103	0·32
B. subtilis, series (i)	(0·090)	283	(0·016)	0·76	156	0·80
B. subtilis, series (ii)	−0·37	252	−0·34	0·65	154	0·59
B. subtilis, series (iii)	(−0·005)	248	−0·16	0·60	169	0·36

clustered about their centroid. This was true also for the bulk of sister-sister pairs; there was, however, a tail to the distribution extending along the bisector of the angle between the axes, indicating very strong association between pairs of sisters of unusually long generation time.

For *Bact. aerogenes* and *Bact. coli anaerogenes* ρ_{MD} is not significantly different from zero, and the immediate conclusion is that delay does not play a large part in determining the variance of generation time. If this is true, neither does it contribute much to the correlation between pairs of sisters (ρ_{SS}) which is yet quite large. There is no constraint implicit in the hypotheses of Kendall and Rahn which requires sister cells to be alike; in fact, independence is assumed. The high value of ρ_{SS} tells heavily either against the hypotheses, or against the relevance to them of the observations. Now ρ_{MD} is compounded of observations made on a number of different occasions. It is to be expected that the true mean value of τ is slightly different for each experiment, and so if ρ_{MD} were really zero a small positive value might be obtained from the combined data. It is possible therefore that the conditions bias both

ρ_{SS} and ρ_{MD} in favour of positive values; ρ_{MD} may be appreciably negative, and ρ_{SS} smaller than it appears to be. However, inspection of the records leaves no doubt as to the great similarity between sister organisms.

In the *Bacillus* group, ρ_{MD} is either zero or negative, but ρ_{SS} remains positive to a high degree of significance. The septate mode of fission appears to be a more highly organized process than the isthmoid mode, and it probably gives rise to very appreciable delay. Its contribution to ρ_{MD} cannot be assessed without knowledge of its variance, though selected examples suggest that the variance may be extremely large, especially in *B. mycoides*. But the differences between the three series of *B. subtilis* indicate that ρ_{MD} is far from being a simple property of the organism. The growth rate in series (ii) and (iii), as will be shown below, was not constant; this should introduce a positive bias, yet ρ_{MD} for (ii) is the most negative of all.

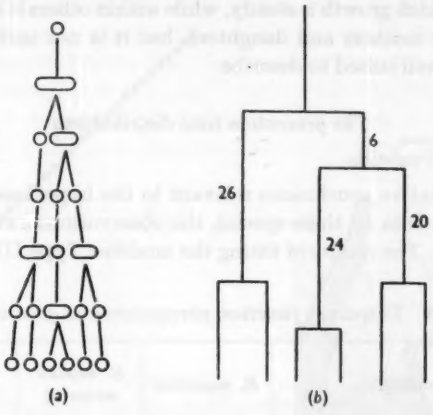


Fig. 3. Abnormal development of *Strep. faecalis*. (*a*) Successive appearance of cells. (*b*) Corresponding family tree; generation times in minutes.

The estimates of ρ_{SS} and ρ_{MD} are affected by the presence of bias in the frequencies. I have attempted a correction in the following way.

(i) If M and D are the generation times of mother and daughter, and $R(M, D)$ the observed frequency of the two together, an estimate of the true frequency is

$$R(M, D)/\alpha(M + D),$$

where α is the biasing factor of equation (2).

(ii) If S_1, S_2 are the generation times of a pair of sisters, and $R(S_1, S_2)$ the frequency of the pair, the corrected value is $R(S_1, S_2)/\alpha\{\max(S_1, S_2)\}$. The ρ calculated from frequencies corrected in this way are also shown in Table 5. No new regularity appears, and the sense of the foregoing remarks is not altered.

Once delayed fission is admitted as a possible contribution to the total variance of generation time, it becomes a matter of importance to assess its variance, since it may turn out so large as to render nugatory any attempt on the present lines to test the hypotheses of Kendall and Rahn. A single overt example of delay was found among the observations on *Strep. faecalis*. Fig. 3*a* shows the appearances; a single cell, instead of dividing when it had reached the usual size, continued to grow until its length was about three times its diameter.

It then split into a sphere and short rod, and fission of the rod followed in 6 min. Fig. 3*b* is the family tree, with the times of fission placed so as to correspond horizontally with Fig. 3*a*. This example is sufficient to engender suspicion that minor degrees of delay may be much more frequent.

Some further progress might be made by a more detailed analysis of correlation among the generation times. I have not attempted such analysis because a direct experimental approach should be possible through the use of ultra-violet or phase-contrast illumination.

I think it likely that the observed values of ρ_{MD} will be found to result from two opposing effects: a small positive correlation, corresponding to the usual succession of τ values, overlaid by less frequent successions of great disparity. Taken together with the large dispersion of within-family variance, this suggests the following picture: a very large family tree, representing the numerical and temporal extension of a culture, consists of a patchwork of areas within some of which growth is steady while within others it is erratic. There is a kind of correlation between mothers and daughters, but it is not such as the crude product-moment coefficient is well suited to describe.

The generation time distributions

(*a*) B. mycoides *and* B. subtilis

Although no quantitative conclusions relevant to the hypotheses of Kendall and Rahn can be drawn from the data on these species, the observations have an important bearing on the theory of fission. The results of fitting the modified Type III and Yule distributions

Table 6. *Frequency function parameters and goodness of fit*

Frequency function		*B. mycoides*	*B. subtilis* series (i)	*B. subtilis* series (ii)	*B. subtilis* series (iii)
Pearson Type III with allowance for bias (eq. (7)):	g	4·07	6·11	7·01	7·47
	m	7·04	3·81	4·21	3·90
	χ^2	19·2	57·3	17·8	15·4
	n	17	15	16	16
	$P(\chi^2)$	0·32	0·000	0·34	0·50
Yule, with allowance for bias (eq. (8)):	g	4·64	9·48	—	—
	m	13·3	8·18	—	—
	χ^2	22·2	74·2	—	—
	n	18	14	—	—
	$P(\chi^2)$	0·22	0·000	—	—

(equations (7) and (8)) are given in Table 6 (see also Fig. 6). Series (i) of *B. subtilis* relates to observations begun at 2½ hr. from inoculation, series (ii) at 4½ hr., series (iii) to observations on organisms growing over diluted sheep blood.

Not only do these species exhibit a higher dispersion of generation time than the rest; the lengths of the organisms are also correspondingly variable, though the parallel between

generation time and length at inception or termination is not exact. The first few vegetative organisms to grow from the spore are often very long—upwards of 30μ—and although the mean length later diminishes, the change is not so well marked as in say *Bact. aerogenes* at the beginning of the logarithmic growth phase. *B. mycoides* cannot be said to possess a normal mean length; the organisms become progressively shorter over a period which may be as long as 24 hr. in some media. While a diminution in mean length is occurring, the number growth rate must be greater than the mass growth rate, and since it is initially less, there is acceleration at some stage. I have therefore checked the uniformity of growth during measurements of generation time in the following way:

It will be recalled that all the experiments grouped into any one column of Table 1 were begun at the same epoch in the life of the culture, and each occupied 80 min. Number growth

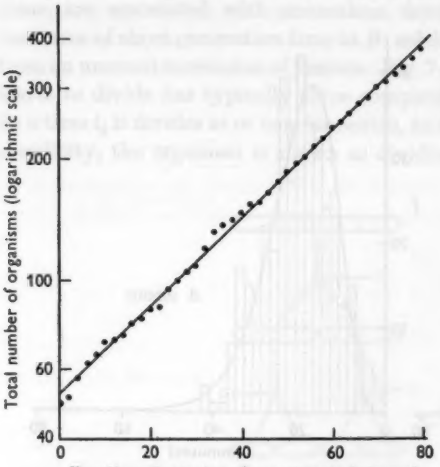


Fig. 4. Number growth rate of *B. mycoides*.

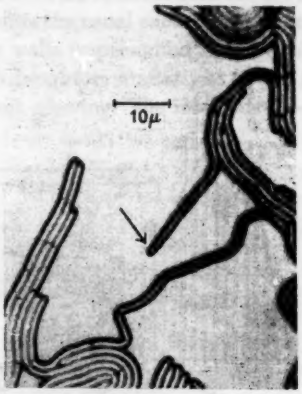


Fig. 5. *B. mycoides* growing on cellophane, 5 hr. from inoculation. A single cell has split off from one end of a long organism.

curves were constructed by adding together the number of organisms under observation at corresponding times in each experiment. For *B. mycoides* and *B. subtilis* series (i) the logarithm of this sum was a linear function of time (Fig. 4), but for *B. subtilis* series (ii) and (iii), there was marked curvature especially over the earlier part of the period. Numerical evaluation of the age distribution by means of equation (4) was also carried out, and compared with the observed age distribution at the 80 min. epoch; series (ii) and (iii) showed an obvious excess of young organisms, as would be expected from an increasing growth rate. These two series are therefore of no quantitative value.

The extraordinarily wide range of generation times in *B. mycoides* (Table 1) is a concomitant of its readiness to divide at any available point (i.e. between any pair of cells). The exaggerated appearance of Fig. 5 is by no means a rarity; the frequency with which a single cell is split off from an organism of eight or more cells is diagnostic of the species. Such single-celled organisms are viable, but develop rather slowly, and some do not divide within 3 hr. of their inception; they are not spores (cf. Bergersen, 1954). *B. subtilis* shows

something of the same freedom in its earlier stages of growth, but it reverts more quickly to regularity. Clearly, in these species the generation time is intimately connected with the coarser structure of the organism, and only remotely with the nuclear fission.

The distribution of *B. subtilis* at $2\frac{1}{2}$–4 hr. (series (i)) is notable for its approximate symmetry (Table 1), in which respect it is quite outstanding; neither the Type III nor Rahn's distribution gives an acceptable fit.

The experiment of growing *B. subtilis* over blood was an attempt to test the constancy of the coefficient of variation of generation time. Pearce and I had found (1951) that in these circumstances the organisms were very short and uniform in size, which suggested that the dispersion would also be small. The blood used in the present work was diluted with saline in order to reduce its viscosity, and over this mixture long organisms were first formed as usual. The mean length gradually diminished, but at the intermediate stages both long and

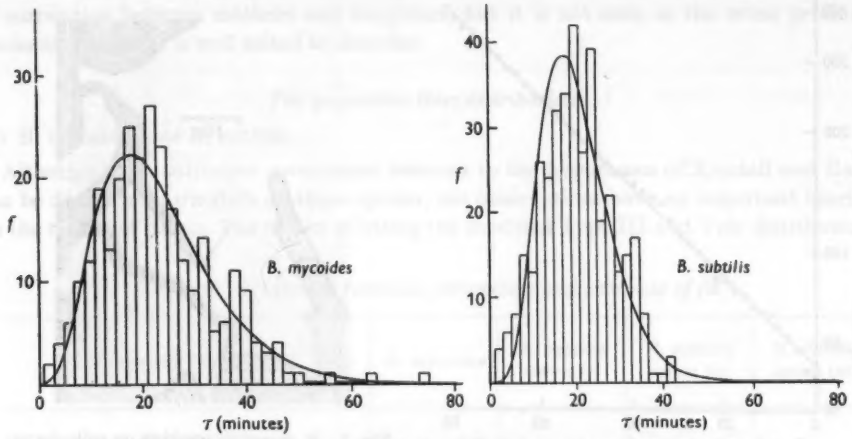


Fig. 6. Generation time distributions of *B. mycoides* and *B. subtilis*, with modified Pearson Type III frequency functions.

short organisms were present, each clone being predominantly of one sort or the other. As I have shown, the growth rate increased considerably during the course of observation. Despite the imperfection of the experimental conditions, the coefficient of variation (0·36) was rather less than in the other two series (0·40 and 0·37).

These rather disconnected remarks indicate that the behaviour of *B. mycoides* and *B. subtilis* is so complex that the experiments are quite insufficient to expose it fully. The distributions, however, possess one regular feature worth examination.

Fig. 6 shows that the frequency of small generation times in *B. mycoides* and *B. subtilis* series (i) is greater than the expectation accorded by the assumed frequency function (a modified Type III). This is true also for *B. subtilis* series (ii) and (iii) (and has since been observed in *B. megatherium*). The observed frequencies near the origin are uniformly so large as to suggest that the true frequency function approaches zero with a non-zero slope (I have pointed out that the observations are less liable to error in this region than elsewhere). In no case did two successive fissions occur in the same 2 min. interval. It will be seen from Fig. 6 that no Pearson Type III or Yule function could give a tolerable fit over the

higher range without being in defect at low values of τ. Both of these functions have a finite positive and non-zero slope at the origin only when $g = 2$ and then

$$\left.\frac{df}{d\tau}\right|_0 = \frac{1}{m^2} \quad \text{and} \quad \frac{2}{m^2} \quad \text{respectively.}$$

But the distributions with $g = 2$ have values of γ_1 and γ_2 quite different from those observed. Kelly & Rahn's results for *B. cereus* are equally suggestive; their grouped frequencies, in order of increasing mean τ, are 9, 2, 4, 5, 17, ...

Among the unicellular organisms, one instance (*Strep. faecalis*, $\tau = 6$ min.) has already been noted of a generation time which was unusually short because of delayed fission in the previous generation. It is perhaps significant that this strain, like the *Bacilli*, exhibited the septate mode of fission. But it cannot be said of *B. mycoides* that the shortest generation times are associated with anomalous development, for reasons already given. Many instances of short generation time in *B. subtilis* can, however, be accounted for as resulting from an unusual succession of fissions. Fig. 7 *a* represents the usual succession; an organism about to divide has typically three completed party walls (represented by pairs of dots). At a time t_0 it divides at or near its centre, and the two daughters similarly at t_{11} and t_{12}. For simplicity, the organism is shown as dividing without growing. The intervals $t_{11} - t_0$ and

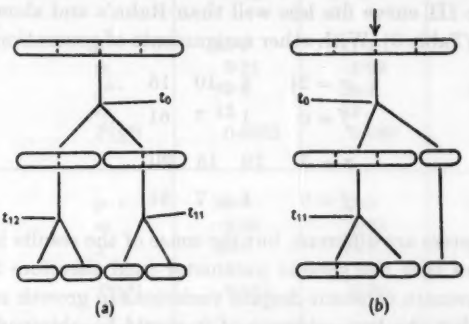


Fig. 7. Normal (*a*) and abnormal (*b*) succession of fissions in *B. subtilis*.

$t_{12} - t_0$ are not generally extreme. Occasionally (Fig. 7 *b*) the first division occurs at a distance of about one-quarter of the length of the organism from one end, and a second division follows very shortly; then the interval $t_{11} - t_0$ is small. That half of the organism to the right of the arrow in Fig. 7 *b* can be usefully regarded as an organism *manqué* with a negative generation time $t_0 - t_{11}$, i.e. the associated separation of cell walls has occurred in the opposite order to the corresponding separation of cytoplasm.

I believe that the impression given by the generation time distributions of *B. mycoides* and *B. subtilis*, namely, that the slope does not vanish near the origin, is a just one, and that the phenomenon is of fundamental significance for the mechanism of fission in *Bacillus*; I hope to deal with it more fully in another communication.

(*b*) Bact. aerogenes, Bact. coli anaerogenes, Strep. faecalis, Pr. vulgaris; *Assessment of the genetic parameter*

The result of fitting the Type III and Yule distributions to the observations on unicellular organisms is shown in Table 7 A–D.

Comparison of sections A and C shows at once that the Type III distribution fits the data better than does Yule's, in each separate case. Moreover, the method of moments gives (section B) substantially the same results as the method of maximum likelihood. Applied to Yule's distribution, it gives a rather better fit, with very different values of the parameters (section D). This disagreement, with the moment fit better than the maximum likelihood, seems to indicate that Yule's distribution is inappropriate to the raw data.

However, the analysis of variance has shown that the crude coefficient of variation (c) of τ must exceed the supposed true and constant value (c_0) because of the experimental dispersion of the mean, and so the method of moments (Table 7 B and D) underestimates g —sampling fluctuations apart. As a further consequence, it does not follow from these results that Kendall's hypothesis is to be preferred to Rahn's.

Kelly & Rahn's results on *Bact. aerogenes* have been shown by Finney & Martin (1951) to be in satisfactory agreement with Rahn's hypothesis so far as the χ^2 test can distinguish. I have already pointed out that there is some doubt as to the correct assignment of mean generation time to the grouped frequencies, but with the same assignment as that adopted by Finney & Martin, viz.

$\tau =$	$2\frac{1}{2}$	$7\frac{1}{2}$	$12\frac{1}{2}$	$17\frac{1}{2}$ ...
$f =$	0	1	7	61 ...

I find that the Type III curve fits less well than Rahn's and shows no advantage in its estimate of c and γ_1 (Table 8). With other assignments of generation times to frequencies

$\tau =$	$2\frac{1}{2}$	5	10	15 ...
$f =$	0	1	7	61 ...

and

$\tau =$	5	10	15	20 ...
$f =$	0	1	7	61 ...

the estimated parameters are different, but the sense of the results is the same.

Hypothesis requires that the genetic parameter (and therefore the true coefficient of variation of τ) shall remain constant despite variations in growth rate. Finney & Martin (1951) pointed out that the best estimate of it would be obtained by fitting frequency functions to the data for each experiment separately, constraining g to be the same for each, but allowing m to differ as required from experiment to experiment. As they realized, this would be an exceedingly laborious process with Rahn's frequency function, but it turns out that for the Type III function a simple maximum-likelihood solution exists.

It is convenient to write the Type III function in the form

$$dF = \frac{g^g \tau^{g-1} e^{-g\tau/a}}{a^g \Gamma(g)}\, d\tau,$$

where the new parameter a is the mean of the distribution, equal to gm. Let a_r be the value of a to be assigned for the rth experiment. Then the likelihood for that experiment is

$$\log L_r = (g-1)\,\Sigma(f \log \tau) - g\Sigma f\tau/a_r - g(\log a_r)\,\Sigma f + g(\log g)\,\Sigma f - \{\log \Gamma(g)\}\,\Sigma f,$$

where the f are the observed frequencies of τ in the rth experiment.

On summing over all experiments and differentiating we obtain the desired maximum-likelihood estimate of g:

$$n\{\psi(\hat{g}) - \log(\hat{g})\} = \sum_r \{\Sigma(f \log \tau)\} - \sum_r \{(\log \hat{a}_r)\,\Sigma f\}, \qquad (9)$$

Table 7. *Frequency function parameters and goodness of fit*

Frequency function and method of fitting		*Bact. aerogenes*	*Bact. coli anaerogenes*	*Strep. faecalis*	*Pr. vulgaris*
(A) Pearson Type III, maximum likelihood:	g	13·4	9·48	12·8	9·08
	m	1·57	2·16	1·95	3·00
	χ^2	16·1	15·1	11·6	27·1
	n	11	12	9	15
	$P(\chi^2)$	0·14	0·23	0·24	0·028
(B) Pearson Type III, moments:	g	13·4	9·53	13·4	9·84
	m	1·57	2·15	1·86	2·77
	χ^2	16·1	15·3	10·2	23·4
	n	11	12	9	15
	$P(\chi^2)$	0·14	0·22	0·33	0·076
(C) Yule, maximum likelihood:	g	32·9	17·1	26·1	14·2
	m	5·21	5·99	6·56	8·48
	χ^2	29·4	24·1	20·0	44·9
	n	12	13	9	16
	$P(\chi^2)$	0·0033	0·030	0·018	0·0002
(D) Yule, moments:	g	59·9	27·2	59·2	29·6
	m	4·50	5·23	5·34	6·84
	χ^2	22·0	13·2	15·9	42·5
	n	10	11	8	14
	$P(\chi^2)$	0·015	0·28	0·043	0·0001
(E) Pearson Type VI, first two moments and σ_x^2:	g	15·1	10·7	15·9	13·1
	s	136	97·7	92·4	45·1
	b	2848	1980	2274	1202
	χ^2	15·4	15·8	11·3	30·1
	n	10	11	8	14
	$P(\chi^2)$	0·12	0·15	0·19	0·0074
(F) Yule-hyperbolic (eq. (14)), first two moments and σ_x^2:	g	79	36	92	55
	l_1	24·3	24·3	29·7	35·1
	l_2	18·1	17·1	20·6	20·6
	χ^2	21·2	14·7	15·1	40·7
	n	10	10	7	13
	$P(\chi^2)$	0·020	0·14	0·035	0·0001

where $\hat{a}_r = \Sigma f\tau/\Sigma f$,

and n is $\sum_r(\Sigma f)$, the total number of observations. The mixed derivatives

$$\frac{\partial^2}{\partial g\,\partial a_r} \quad \text{and} \quad \frac{\partial^2}{\partial a_r\,\partial a_s} \quad (r \neq s)$$

of $\sum_r \log L_r$ are all zero (this is the virtue of using the parameter a instead of m), and so, as in the simple case, the asymptotic variance of $\hat{g}$ can be estimated as

$$\operatorname{var}\hat{g} = 1/n\{\Gamma_2(\hat{g}) - 1/\hat{g}\}.$$

Table 8. *Comparison of Pearson Type III and Yule distributions applied to Kelly & Rahn's results for* Bact. aerogenes

	g	m	χ^2	n	$P(\chi^2)$	c	γ_1
Type III	10·8	2·90	17·0	8	0·030	0·30	0·61
Yule	26·0	8·14	14·9	9	0·094	0·33	1·18
From data	—	—	—	—	—	0·31	0·91

The results of applying equation (9) to the data for the unicellular organisms are shown in Table 9 E. The g values obviously differ significantly from species to species, and all are less than 20, the approximate figure arrived at by Kendall (1948) from a consideration of Kelly & Rahn's experiments on *Bact. aerogenes*. In spite of the fact that the variance of τ in the present measurements on that organism is less than in theirs, the value of g is still only 16·5; Kendall, of course, was not aware that the apparent heterogeneity of the data was largely intrinsic.

Although the preceding form of treatment is not available for Rahn's hypothesis, a satisfactory approach can in any case be made through the coefficient of variation.

A number of generation time experiments carried out on one species can be considered to provide a sample of frequency functions $h(\tau, a, c_0)$ drawn from a population in which the true coefficient of variation $c_0 = \sigma/a$ is by hypothesis constant, but the mean a is dispersed. In this hypothetical population of h-functions, let $j(a)$ be the frequency function of the mean. Then considering the whole series of experiments together, the expectation of an observation τ will be

$$dF = \int_{a_l}^{a_u} h(\tau, a, c_0)\, j(a)\, da\, d\tau, \tag{10}$$

where a_u and a_l are the limits within which a must lie. The moments about zero are then

$$\mu'_r = \int_0^\infty \int_{a_l}^{a_u} \tau^r h(\tau, a, c_0)\, j(a)\, da\, d\tau.$$

It can safely be assumed that the order of integration may be inverted, so that

$$\mu'_r = \int_{a_l}^{a_u} \int_0^\infty j(a)\, \tau^r h(\tau, a, c_0)\, d\tau\, da.$$

In particular,

$$\mu_1' = \int_{a_l}^{a_u} aj(a)\,da = \mathbf{a}, \quad \text{say.}$$

$$\mu_2' = \int_{a_l}^{a_u} j(a)\,a^2(c_0^2+1)\,da = (c_0^2+1)\,\mu_2'(j), \quad \text{say.}$$

Hence the observed crude coefficient of variation, c, is related to c_0 by

$$c^2 = \mu_2'/(\mu_1')^2 - 1 = (c_0^2+1)\,\mu_2'(j)/\mathbf{a}^2 - 1.$$

(Evidently $c \geqslant c_0$ always, since $\mu_2'(j) \geqslant \mathbf{a}^2$.) Or, more symmetrically, writing c_j for the coefficient of variation of the j-distribution,

$$c^2+1 = (c_0^2+1)\,(c_j^2+1). \tag{11}$$

Thus from a crude coefficient c a corrected value c_0 can be calculated if the experimental variance of the mean is known, always under the hypothesis that c_0 is constant. The necessary figures are already available from the working of the analysis of variance; the σ_E^2 of Table 3 are estimates of $\mu_2(j)$ and

$$c_j^2 = \mu_2(j)/\mathbf{a}^2.$$

Application of equation (11) then yields the figures in Table 9 B.

It is to be noted that the relation between g and c_0 so calculated is determined by the h-distribution, and is independent of $j(a)$.

If from the corrected values c_0 new estimates of Kendall's $g\,(=1/c_0^2)$ are calculated, the results agree well with the revised maximum likelihood estimates (Table 9 C and E), and so the assumption that the h-distribution is of Type III is not an unreasonable one.

It is now possible to derive also improved values of Rahn's g from

$$c_0^2 = \frac{\Gamma_2(1) - \Gamma_2(g+1)}{\{\Gamma_1(g+1) - \Gamma_1(1)\}^2}.$$

These values are given in Table 9 D.

Equation (10) suggests a method of developing modified frequency distributions capable of representing the crude data more accurately, and of being subjected to a meaningful significance test. Each series of experiments furnishes a number of values a_r and a variance σ_E^2. The exact distribution of a is not known, but a graphical examination of the individual a_r shows a marked modal tendency. Suppose then, for convenience, that a is distributed as a Pearson Type V variate:

$$j(a) = \frac{b^s a^{-s-1} e^{-b/a}}{\Gamma(s)},$$

with overall $\bar{\tau} = \mathbf{a} = b/(s-1)$,

and $\sigma_E^2 = \mu_2(j) = b^2/(s-1)^2\,(s-2)$.

Then if the distribution $h(\tau, a, c_0)$ is of Type III, equation (10) is easily integrated, and becomes

$$dF = \frac{(b/g)^s}{B(g,s)}\,\frac{\tau^{g-1}}{(\tau + b/g)^{g+s}}\,d\tau, \tag{12}$$

where $B(g,s)$ is the complete beta-function. This is a Pearson Type VI distribution.

I have fitted equation (12) to the raw data by using the first two moments and the variance σ_E^2 of the mean, to determine b, s and g (Table 7 E; the values of g are of course the

same as those in Table 9 C, since the j-distribution is irrelevant). To neglect σ_E^2 and use the first three moments would omit part of the information available, viz. the knowledge of how the τ were grouped by experiments. The goodness of fit is no better than that of the simple Type III function. One of the fitted curves—*Pr. vulgaris*, for which σ_E^2 is largest—is shown in Fig. 8 (curve (i)).

Table 9. *The crude* (c) *and corrected* (c_0) *coefficients of variation, with final estimates of the genetic parameter* (g)

	Bact. aerogenes	*Bact. coli anaerogenes*	*Strep. faecalis*	*Pr. vulgaris*
(A) c	0·273	0·324	0·273	0·319
(B) c_0	0·257	0·306	0·251	0·277
(C) g (Kendall) from c_0	15·1	10·7	15·9	13·1
(D) g (Rahn) from c_0	79	36	92	55
(E) g (Kendall) by maximum likelihood	16·5 ± 1·1	10·9 ± 0·7	18·6 ± 2·3	14·4 ± 1·0

Another convenient choice of $j(a)$ is to take it as rectangular with mean $\mathbf{a}$ and range $2\sqrt{3}\,\sigma_E$. Then, still with a Type III distribution for $h(\tau, a, c_0)$, equation (10) becomes

$$dF = \frac{g}{2\sqrt{3}\,\sigma_E(g-1)}\left\{I\left(\frac{\tau g}{\mathbf{a}+\sqrt{3}\,\sigma_E}, g-1\right) - I\left(\frac{\tau g}{\mathbf{a}-\sqrt{3}\,\sigma_E}, g-1\right)\right\} d\tau, \tag{13}$$

where $I(x, p)$ is the incomplete gamma function

$$\frac{1}{\Gamma(p)}\int_x^\infty \xi^{p-1} e^{-\xi}\, d\xi.$$

This distribution, fitted to the *Pr. vulgaris* data, is also shown in Fig. 8 (ii). The χ^2 is 29·5, as against 30·1 for the Type VI, and the close geometrical similarity of the two curves shows that it matters little what functional form is chosen for $j(a)$. *A fortiori* this is true for the other organisms, for which σ_E^2 is much smaller.

Yule's distribution does not combine readily, in equation (10), with any of the standard distributions. However, after the result of the previous paragraph, we may take

$$j(a) = \frac{\lambda}{a} \quad (l_1 \geqslant a \geqslant l_2),$$

i.e. $j(a)$ is a segment of a rectangular hyperbola. Then

$$\lambda = 1/\log(l_1/l_2),$$
$$\mathbf{a} = \lambda(l_1 - l_2),$$
$$\sigma_E^2 + \mathbf{a}^2 = \mu_2'(j) = \lambda(l_1^2 - l_2^2).$$

Equation (10) becomes

$$dF = \int_{l_2}^{l_1} \frac{\lambda g \Psi}{a^2} e^{-\tau\Psi/a}(1 - e^{-\tau\Psi/a})^{g-1} da\, d\tau,$$

where $\Psi = \Gamma_1(g+1) - \Gamma_1(1)$; by equation (5), $a = m\Psi$, and by hypothesis Ψ is constant. Integration is immediate on writing $x = e^{-\tau\Psi/a}$:

$$dF = \frac{\lambda}{\tau}[\{1-e^{-\tau\Psi/a}\}^g - \{1-e^{-\tau\Psi/a_1}\}^g]\,d\tau. \tag{14}$$

Like the Type VI, this distribution fits no better than the parent Yule distribution with smaller g (Table 7 D and F).

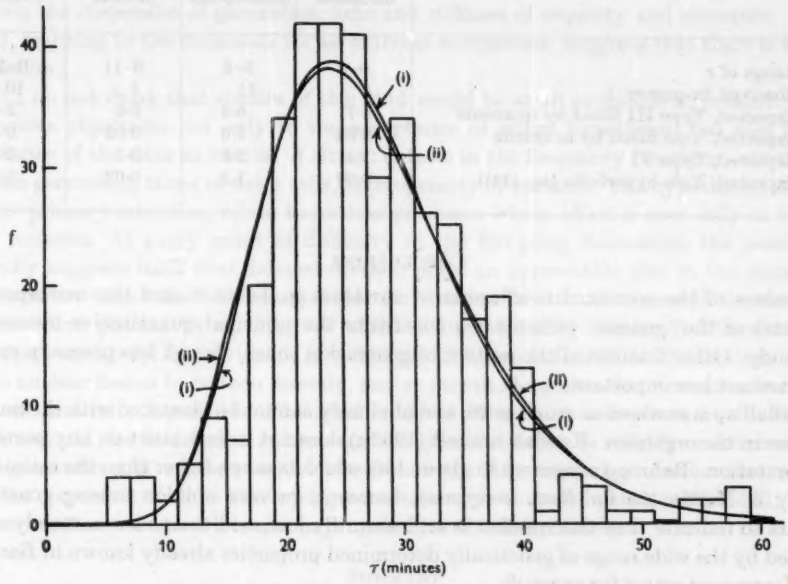


Fig. 8. Generation time distribution of *Pr. vulgaris* with fitted curves.
(i) Pearson Type VI, (ii) distribution of equation (13).

As in the *Bacillus* group, there is in all four species an excess of observed over expected frequencies near the origin (Table 10); this is true for all the distributions of Table 7. The organisms of very short generation time are well distributed among the several experiments carried out on each species; thus the ten *Pr. vulgaris* organisms of $\tau \leqslant 10$ (see Fig. 8) are distributed among five experiments—they are not associated with a single experiment of exceptionally low mean τ. And it is evident that if the restriction, c_0 = constant, is maintained, only extravagant dispersion of the mean could appreciably increase the ordinates of the fitted curves near the origin. Qualitatively, this is just such an effect as might result from the occurrence of occasional large delays in fission; because of the skewness of the distribution the presence of a short-lived daughter organism would be much more obvious than that of its long-lived mother.

On the score of goodness of fit, therefore, Kendall's hypothesis is to be preferred to Rahn's, but the proposed frequency functions do not represent the observations altogether satisfactorily. The large scatter of within-family variance (Table 2) is admissible only on Rahn's

hypothesis, but it is associated with the presence of an excess of organisms of very short generation time, and cannot be solely accounted for by the long upper tail of the Yule distribution.

Table 10. *Observed and expected frequencies of short generation times*

(The range of τ is, for uniformity, taken to be half the distance of the mode from the origin.)

	Bact. aerogenes	*Bact. coli anaerogenes*	*Strep. faecalis*	*Pr. vulgaris*
Range of τ	0–9	0–9	0–11	0–11
Observed frequency	5	11	1	10
Expected, Type III fitted by moments	1·7	6·8	0·6	3·4
Expected, Yule fitted by moments	0·04	1·6	0·02	0·4
Expected, Type VI	1·4	5·9	0·5	2·2
Expected, Yule-hyperbolic (eq. (14))	0·03	1·2	0·02	0·3

Conclusions

The values of the corrected coefficients of variation in Table 9, and the corresponding estimates of the 'genetic' parameter g constitute the principal quantitative outcome of this study. Other features of the pattern of generation times, though less precisely expressible, are not less important.

Kendall's g is nowhere as much as 20, and obviously cannot be identified with the number of genes in the organism. Kendall himself (1952*a*) does not in fact insist on any particular interpretation. Rahn's g ranges up to about 100, which is much larger than the estimate of Finney & Martin (25 for *Bact. aerogenes*); however, private opinion among geneticists appears to indicate that this number is still absurdly low, and their view certainly seems justified by the wide range of genetically determined properties already known in *Bact. coli* and *Neurospora crassa* for example.

The observations on *B. mycoides* and *B. subtilis*, though possessing a specialized interest of their own, are not at present susceptible of any simple interpretation, however tentative.

It is very likely that a mechanism of Kendall's type, i.e. a stepwise process, does occur during the fission of an organism, and there is no *a priori* reason why it should not also be preceded or accompanied by Rahn's gene-duplication process. Then the one will be manifested clearly only if it is slow enough, relative to the other, to be the dominant factor in determining generation time. If neither is dominant (and this is a real possibility) the experimental picture will be confused, though it may not be recognized to be so. As things stand, Kendall's hypothesis is to be preferred, but only on the ground that its formal expression as a Pearson Type III distribution is in fair accord with the data. Kendall and Waugh (Kendall, 1952*a*) have also proposed a modification of Kendall's original hypothesis by relaxing the condition that the duration of the primitive steps of his fission process shall have the same frequency function for all. The modified hypothesis subsumes Rahn's as a special case (though only formally, not structurally). I have not examined its practical consequences in detail, but it has one significant feature pointed out by Kendall: suppose g is estimated by fitting a Type III distribution to a set of data, then if the primitive steps are not all distributed alike, their number will be greater than g.

It would obviously be possible to test Rahn's hypothesis by measuring generation times of the same species at different temperatures and when growing on different media; the coefficient of variation should remain constant. Under Kendall's more reserved hypothesis, change of temperature would be expected to make no difference, but his g might well depend on the number of synthetic processes demanded by the laying down of nuclear matter or cell wall, for instance, and so might depend in turn on the chemical complexity of the nutrients offered. Further, a similar study of a wide range of organisms, including autotrophs and the most exigent pathogens, would show whether or not there is any correlation between the dispersion of generation time and richness of capacity and structure. Rahn (1932), reaching to the mammals for an extreme comparison, suggests that there is such a correlation.

But I do not think that studies of this kind would be at all profitable at present. There are serious objections not only to the acceptance of either hypothesis, but also to the acceptance of the data as critical of either: defects in the frequency functions, correlation between generation times of sister cells, heterogeneity of variance. The hypotheses, at least in their primary intention, relate to nuclear processes whose effect is seen only at two or three removes. At every point of difficulty in the foregoing discussions the possibility naturally suggests itself that delayed division plays an appreciable part in the dispersion of generation time; that the recognizable termination of the cell succeeds the essential determinative process by an interval which is itself sensibly dispersed. Therefore, the immediate need is for improvement in technique. The use of ultraviolet illumination should enable nuclear fission to be seen directly, and so permit the generation time so-called to be analysed into its components. Extended study of cultures on cellophane over a flowing medium would also be valuable; when constant growth rate can be reliably maintained over many generations, it will be possible to calculate the dispersion of generation time simply from the dispersion of clone size, with great economy in time and patience (Kendall, 1948). Further mathematical work is also in progress (see Kendall, 1952*b*).

SUMMARY

1. The conditions are discussed which must be met in any attempt to measure a frequency distribution of generation times of micro-organisms.

2. Generation times of individual organisms of six species have been measured: *Bacterium aerogenes*, *Bact. coli anaerogenes*, *Streptococcus faecalis*, *Proteus vulgaris*, *Bacillus subtilis*, *B. mycoides*.

3. The hypotheses developed by Kendall and by Rahn each connect dispersion of generation time with a postulated mechanism of fission; each implies a definite mathematical form for the distribution. Comparison with experiment suggests that Kendall's is to be preferred.

4. In Rahn's hypothesis, one of the parameters of the distribution is identified with the number of genes in the organism. When every allowance is made for experimental error, and in the most favourable case (*Strep. faecalis*), the estimated number of genes is less than 100.

5. Generation times are not distributed at random, but are in part determined by a weak and variable hereditary factor effective only over a few generations.

6. There are positive objections to the acceptance of either Kendall's or Rahn's hypothesis: correlation between generation times of sister cells, heterogeneity of variance in

small families of organisms, certain defects in the proposed distribution functions. These difficulties, and the effect mentioned under (5), may be due to delayed fission, that is, to the lapse of an appreciable period between the division of the nucleus and the observed separation of the cell into two parts.

7. In the *Bacillus* group, the measurements are not relevant to the hypotheses, because the organisms are multicellular.

8. Improved techniques are required for the further pursuit of the subject.

T. W. Pearce and Jean M. Scott have shared with me the labour of observation. Much of the arithmetic was carried out by R. Ash. I am indebted to Dr D. W. Henderson for his confidence and encouragement and to S. Peto for his helpful criticism. In the analysis of data I have benefited greatly from Prof. E. S. Pearson's guidance. Publication is by permission of the Chief Scientist, Ministry of Supply.

REFERENCES

BARTLETT, M. S. (1937). Properties of sufficiency and statistical tests. *Proc. Roy. Soc.* A, **160**, 268–82.

DE BARY, H. A. (1884). *Vergleichende Morphologie und Biologie der Pilze, Mycetozoen, und Bakterien.* Leipzig.

BERGERSEN, F. J. (1954). The filamentous forms of *Bacillus megaterium*. *J. Gen. Microbiol.* **11**, 175–9.

BISSET, K. A. (1950). *The Cytology and Life-history of Bacteria.* Edinburgh: E. and S. Livingstone.

BOX, G. E. P. (1953). Non-normality and tests on variances. *Biometrika*, **40**, 318–35.

DANIELS, M. E. (1952). The covering circle of a sample from a circular normal distribution. *Biometrika*, **39**, 137–43.

DAWSON, I. M. & STERN, H. (1954). Changes in the bacterial cell wall during cell division as seen in the electron microscope. *J. Gen. Microbiol.* **10**, iii.

FINNEY, D. J. & MARTIN, L. (1951). A re-examination of Rahn's data on the number of genes in bacteria. *Biometrics*, **7**, 133–44.

FISHER, R. A. (1922). On the mathematical foundation of theoretical statistics. *Philos. Trans.* A, **222**, 309–68.

HARRIS, NORA K. & POWELL, E. O. (1951). A culture chamber for the microscopical study of living bacteria, with some observations on the spore-bearing aerobes. *J. R. Micr. Soc.* **71**, 407–20.

HARRIS, T. E. (1951). Some mathematical models for branching processes. *Proc. Second Berkeley Symposium on Mathematical Statistics and Probability*, p. 305.

KELLY, C. D. & RAHN, O. (1932). The growth rate of individual bacterial cells. *J. Bact.* **23**, 147–53.

KENDALL, D. G. (1948). On the role of a variable generation time in the development of a stochastic birth process. *Biometrika*, **35**, 316–30.

KENDALL, D. G. (1952*a*). On the choice of a mathematical model to represent normal bacterial growth. *J. R. Statist. Soc.* B, **14**, 41–4.

KENDALL, D. G. (1952*b*). Les processus stochastiques de croissance en biologie. *Ann. Inst. Poincaré*, **13**, 43–108.

LOMINSKI, I. & LENDRUM, A. C. (1947). The mechanism of swarming of *Proteus*. *J. Path. Bact.* **59**, 688–91.

ORSKOV, J. (1922). Method for the isolation of bacteria in pure culture from single cells and procedure for the direct tracing of bacterial growth on a solid medium. *J. Bact.* **7**, 537–49.

PEARCE, T. W. & POWELL, E. O. (1951). New techniques for the study of growing micro-organisms. *J. Gen. Microbiol.* **5**, 91–103.

RAHN, O. (1932). A chemical explanation of the variability of the growth rate. *J. Gen. Physiol.* **15**, 257–77.

SCHAUDINN, F. (1902). Beiträge zur Kenntnis der Bakterien und verwandter Organismen. I. *Bacillus bütschlii*, n.sp. *Arch. Protistenk.* **1**, 306–43.

SHERMAN, J. & EWELL, R. B. (1942). A six-place table of Einstein functions. *J. Phys. Chem.* **46**, 641–61.

YULE, G. U. (1925). A mathematical theory of evolution, based on the conclusions of Dr J. C. Willis, F.R.S. *Philos. Trans.* B, **213**, 21–87.

QUANTUM HYPOTHESES

By S. R. BROADBENT

The British Coal Utilisation Research Association

1. Introduction

(1·1) Statistical problems which involve a mixed population are complex, even when the component subpopulations can be assumed to be normal. Karl Pearson (1894) noted in his dissection of a frequency curve into two normal components, 'It may happen that we have a mixture of $2, 3, \ldots, n$ homogeneous groups, each of which deviates about its own mean symmetrically, and in a manner represented with sufficient accuracy by the normal curve.... The equations for the dissection of a frequency curve into n normal curves can be written down as for the special case of $n = 2$ treated in this paper; they require us only to calculate higher moments. But the analytical difficulties, even for the case of $n = 2$, are so formidable, that it may be questioned whether the general theory could ever be applied in practice to any numerical case.'

The present paper discusses the simpler problem in which the means of the components are equally spaced. The hypothesis that this is the case has been called by Hammersley & Morton (1954) a *quantum hypothesis*, for the means are then a constant plus multiples of a basic quantity or quantum.

Examples of such distributions will be found in the paper and in the references cited. It is worth noting that such rules as Brook's law (1886), formulated by Fowler (1909) in his study of the Ostracoda, that 'during early growth, each stage increases at each moult by a fixed percentage of its length, which is approximately constant for its species and sex', and Przibram's rule (1912) discussed by Wigglesworth (1942), that 'the weight doubles at each instar, and at each moult all linear dimensions are multiplied by $\sqrt[3]{2}$', take the form of quantum hypotheses when the logarithms of the weights or lengths are taken. Therefore situations in which a quantum hypothesis applies after the variate is transformed are here considered.

(1·2) Data suggest a quantum hypothesis by the occurrence of regularly spaced modes. From such data three results are commonly required:

(i) an estimate of the quantum which determines the spacing of the modes,

(ii) an estimate of the scatter within the subdistributions, and

(iii) a demonstration that the quantum rule is not disobeyed, i.e. that the data are not more likely to come from some other distribution, perhaps unimodal.

When the valleys between the modes are very well marked, there is little difficulty in meeting these requirements. Controversy about the interpretation of some data shows that a statistical treatment is required for those cases in which grouping at the modes is less obvious.

In the most difficult situation we are presented with data alleged to support a quantum hypothesis. No independent information is available; the hypothesis has arisen from a study of the data and the means are estimated from the observations. We are required to test whether the observations are genuinely grouped about these means. For examples of this type, see Hammersley & Morton's (1954) Druid Circle problem and Grant's (1952) measure-

ments of the energy levels of atomic nuclei. Hammersley & Morton conjecture that in this form the problem is beyond present-day analytic resolution, and offer a Monte Carlo method of testing the supposed grouping. In this case the difficulty is that the data which prompted the hypothesis are also used for estimation and for testing, and it is not clear what allowance can be made in a test for such previous use. The warning given by Pearson & Chandrasekar (1936) is relevant: 'To base the choice of the test of a statistical hypothesis upon an inspection of the observations is a dangerous practice; a study of the configuration of the sample is almost certain to reveal some feature, or features, which are exceptional if the hypothesis is true....By choosing the feature most unfavourable to the hypothesis out of a very large number of features examined, it will usually be possible to find some reason for rejecting the hypothesis.' We need add only that it will be equally possible to choose some untrue hypothesis and to find some test which favours it.

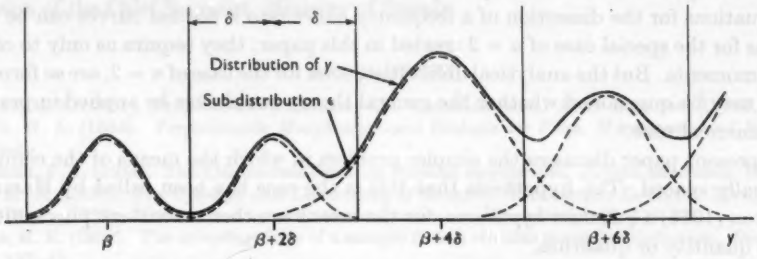


Fig. 1. Frequency curve on the quantum hypothesis.

(1·3) The population we shall consider will be compounded of normal subdistributions with means equally spaced at $\beta+2r\delta$ $(r=0,1,\ldots,m)$. The probability that a randomly chosen member of the population is from the tth subdistribution, i.e. has expectation $\beta+2t\delta$, is p_t; not all the p_t are supposed non-zero, but $\sum_{t=0}^{m} p_t = 1$. The frequency function of such a population is shown in Fig. 1. The subdistributions may be homoscedastic, with common variance σ^2, or the S.D. may increase linearly with the mean.

(1·4) The frequency histogram of n observations from such a population will usually consist of regularly spaced peaks if σ/δ is small and the p_t are not too different. As σ/δ increases the valleys between the peaks disappear and it becomes difficult to assign an observation to its correct subdistribution. In these circumstances the p_t determine the overall appearance of the population, i.e. whether it is unimodal, symmetrical and so on. It is clear that using this model and by suitable choice of the parameters we can approximate with any required precision to any distribution, and that the graduation between multimodal and unimodal distributions is continuous.

Just as the genuine quantum model merges into the unimodal model, so it is possible to find a quantum which fits data from any distribution, since experimental data are usually rational. We have only to take the H.C.F. of the observations, after adding an arbitrary constant, to obtain a quantum which fits the data precisely. Other constants are possible which fit the data less exactly, and even if the observations were not rational, a quantum can be found to fit the data as closely as we please. Equally, when a quantum fits the data we have only Occam's principle to exclude the possibility that the true quantum is a half

or a third of the one we have accepted. The difficulties of estimation in a similar situation have been discussed by Hammersley (1950).

Obviously when so many alternatives are so easily produced we must be careful to state explicitly what assumptions are being made and where independent evidence is being used.

(1·5) In this paper the problems considered will be limited to the following:

(i) Estimation of the positions of the modes when the observations have been allotted to the correct subdistributions, and these subdistributions are normal.

(ii) Estimation of the variance of each component when the subdistributions are normal and *either* the observations have been allotted to the correct subdistributions (positions of the modes unknown) *or* the positions of the modes are known (observations not allotted to subdistributions).

(iii) Testing whether to accept a hypothesis which specifies the positions of the modes and which is independent of the data used in the test. The normality of the subdistributions is not here assumed, the p_t are not specified, nor are the observations allotted to subdistributions. The hypothesis against which the quantum hypothesis is compared is one of a class described in § (4·2).

2. ESTIMATION: MODES

(2·1) In this section it is assumed that the data can be allotted to the correct subdistributions, i.e. that when we are given any observation we can say it comes from the tth subdistribution, although we do not know precisely the location of this subdistribution. This may be possible because the data fall into clearly defined groups with every component known to be represented or for some other reason. For example, in the measurement of insects during moulting, each group or subdistribution is composed of measurements after a defined number of moults.

Let y_{rs} be the sth observation in the rth subdivision, with mean $\beta + 2r\delta$. Then

$$y_{rs} = \beta + 2r\delta + \epsilon_{rs},$$

where β and δ are unknown constants (2δ is the quantum), r is zero or a positive integer, and ϵ_{rs} the normal error, or deviation of y_{rs} from its mean, with mean zero. Let $r = 0, 1, \ldots, m$; $s = 1, 2, \ldots, n_r$; $\sum_{r=0}^{m} n_r = n$; $\sum_{s=1}^{n_r} y_{rs} = Y_r$.

(2·2) The method that has frequently been used to detect the quantum situation and to estimate the positions of the modes is to plot the means of the successive groups (or the experimental modes; these may also be used when observations cannot be allotted to groups) against r, and to draw a straight line through the points so obtained. A similar but efficient estimation of β and 2δ is by the regression of y_{rs} on r. The method is equivalent to least-squares and to maximum-likelihood procedures for estimating the means of the subdistributions subject to the restrictions that they are in arithmetic progression. The results of the regression analysis are summarized in the three sections below.

(2·3) ϵ_{rs} *has* S.D. σ (*all* r). The equations giving estimates b and $2d$ for β and 2δ are

$$b = (\Sigma r^2 n_r \Sigma Y_r - \Sigma r n_r \Sigma r Y_r)/\Delta,$$

$$2d = (n\Sigma r Y_r - \Sigma r n_r \Sigma Y_r)/\Delta,$$

where $\Delta = n\Sigma r^2 n_r - (\Sigma r n_r)^2$, and all summations are over $r = 0, 1, \ldots, m$.

The properties deducible from regression coefficients apply to b (the intercept of the regression line with the ordinate) and $2d$ (the slope). We have:

(i) the variance σ^2 of each y_{rs} about its mean is estimated by $s_1^2 = S_1^2/(n-2)$, where

$$S_1^2 = \{\Delta\Delta' - (n\sum_r rY_r - \sum_r rn_r \sum_r Y_r)^2\}/(n\Delta),$$

and

$$\Delta' = n\sum_r \sum_s y_{rs}^2 - (\sum_r Y_r)^2.$$

(ii) S_1^2 has σ^2 times the χ^2 distribution with $(n-2)$ d.f.

(iii) the variance of $2d$, estimator of the quantum 2δ, is itself estimated by s_1^2/Δ.

(iv) s_1^2/Δ is independent of $2d$, so the t-test may be applied to hypotheses about 2δ.

(v) The variance of b, estimator of β, is itself estimated (not independently of the variance of $2d$) by

$$s_1^2(1 + (\sum_r rn_r)^2/n\Delta)/n.$$

(2·4) $\beta = 0$ *given;* ϵ_{rs} *has* S.D. σ *(all* r*).* The estimator for 2δ is

$$2d = (\sum_r rY_r)/(\sum_r r^2 n_r).$$

(i) the variance σ^2 of each y_{rs} about its mean is now estimated by $s_2^2 = S_2^2/(n-1)$, where

$$S_2^2 = \sum_r \sum_s y_{rs}^2 - (\sum_r rY_r)^2/\sum_r r^2 n_r.$$

(ii) S_2^2 has σ^2 times the χ^2 distribution with $(n-1)$ d.f.

(iii) the variance of $2d$ is itself estimated by $s_2^2/\sum_r r^2 n_r$.

(iv) $s_2^2/\sum_r r^2 n_r$ is independent of $2d$, so the t-test is applicable to hypotheses about 2δ.

(2·5) $\beta = 0$ *given;* ϵ_{rs} *has* S.D. $r\sigma$. This is the case in which the S.D. of each observation varies directly with its expected value, i.e. with $2r\delta$ and so with r. In the least-squares equation for estimating 2δ each observation must be weighted inversely by its S.D.; we then obtain the maximum-likelihood estimate

$$2d = \sum_r (Y_r/r)/n.$$

This is the arithmetic mean of the homoscedastic ratios Y_r/r, i.e. in the regression model each observed point gives an equally precise estimate of the slope.

The variance σ^2 is estimated by essentially the same method as that implied in (2·3) and (2·4). Consider the identity

$$(y_{rs} - 2r\delta)/r = (y_{rs} - 2rd)/r + 2(d-\delta).$$

Of these quantities the first is distributed normally with mean zero and variance σ^2, the second does not need δ in its calculation, while the third has mean zero and variance σ^2/n. It follows that

$$\sum_r \sum_s (y_{rs} - 2r\delta)^2/r^2 = \sum_r \sum_s (y_{rs} - 2rd)^2/r^2 + 4n(d-\delta)^2,$$

and by Cochran's theorem that

$$S_3^2 = \sum_r \sum_s (y_{rs} - 2rd)^2/r^2$$

has σ^2 times the χ^2 distribution with $(n-1)$ d.f.

(i) The variance σ^2 is estimated by $s_3^2 = S_3^2/(n-1)$, where S_3^2 is calculated from

$$S_3^2 = \sum_r \sum_s (y_{rs}/r)^2 - \{\sum_r (Y_r/r)\}^2/n.$$

This is otherwise clear, since y_{rs}/r has variance σ^2.

(ii) The variance of $2d$, estimator of 2δ, is itself estimated by s_3^2/n.

(iii) s_3^2 is independent of $2d$, so the t-test may be applied to hypotheses about 2δ.

(iv) If we wish to set limits to a single new observation in the rth group, we must take into account the uncertainty of the supposed mean $2rd$, which contributes a variance $r^2\sigma^2/n$, and the variance $r^2\sigma^2$ of a single observation in this group. The limits are

$$r\{2d + t_{n-1}s_3\sqrt{[(n+1)/n]}\},$$

where t_{n-1} denotes the value of Student's t with $(n-1)$ d.f. at the appropriate probability level.

(2·6) *Example of estimation.* Svedberg (1939) gives the molecular weights of fifty-six proteins determined by sedimentation velocity or by sedimentation equilibrium. For twenty proteins both methods have been used so that seventy-six measurements in all are given. Svedberg deduces a 'law of simple multiples....If we choose 17,600 as the unit the majority of the proteins may be divided into eleven classes with molecular weights which are multiples of this unit by factors containing powers of 2 and 3. The rule is only approximate, indicating that the underlying principle is obscured by some secondary factor.' He notes seventy observations in eleven classes which he considers obey this law, and he gives the factor relevant to each class by which the unit is multiplied.

We take these seventy observations as a random sample from a population of the type considered in § (2·5), and for each class take the factor given by Svedberg as r. It is clear from a study of the data that it is appropriate to suppose the S.D. of each class is linearly related to its mean. We can then apply the method of § (2·5) to estimate the quantum for which Svedberg gives the value 17,600. We obtain $2d = 17{,}920$, $s_3^2 = 3{\cdot}83$ (69 d.f.), i.e. 95 % limits to 2δ are 17,460 − 18,390.

Six observations are so far from Svedberg's supposed modes that he did not classify them. When the limits within which a new observation may be expected to lie are calculated it is found that two of the outliers are within 95 % limits, two within 98 % limits, and two within 99 % limits. Therefore these extreme observations are not inconsistent with the remainder of the data.

It may be remarked here that Johnston, Longuet-Higgins & Ogston (1945) consider that Svedberg's data do not support his hypothesis. Their tests will be discussed below in § (4·6).

3. ESTIMATION: VARIANCE

(3·1) In §§ (2·3), (2·4) and (2·5) an estimator for σ^2 and the form of its distribution were obtained when the positions of the modes were estimated.

(3·2) Now suppose the positions of the modes are known, independently of the data, but that the observations have not been allotted to subdistributions and may overlap, i.e. lie nearer to another mean than their own. The weights p_i are not known. The distribution is that of § (2·1).

Let z_i $(i = 1, 2, \ldots, n)$ be the distance of y_{rs} from the nearest mode, i.e.

$$z_i = y_{rs} - (\beta + 2r'\delta),$$

where r' is chosen to minimize $|z_i|$. Clearly $|z_i| \leqslant \delta$, and z_i is ϵ_{rs} plus or minus an integer (or zero) multiple of 2δ. For those ϵ_{rs} satisfying $|\epsilon_{rs}| \leqslant \delta$, the distribution of z_i is the same as

that of ϵ_{rs}, i.e. normal with mean zero and variance σ^2 truncated at $\pm\delta$. For those ϵ_{rs} satisfying $\delta < \epsilon_{rs} \leqslant 3\delta$, the distribution of z_i is the same as that of $(\epsilon_{rs} - 2\delta)$ truncated at $\pm\delta$, and so on. The effect of transforming from y to z is to cut the overall frequency distribution of y at $\ldots, \beta-\delta, \beta+\delta, \beta+3\delta, \ldots$ and to lump together the truncated portions. Since each subdistribution is the same, save for its mean and p_i, the result is shown in Fig. 2, where the subdistribution (dotted line) is transformed into the distribution of z (full line) by successively 'turning in' the tails of the distribution. It will be noticed that the p_i do not appear in the distribution of z, nor is the allocation of y_{rs} to its correct subdistribution relevant to z_i.

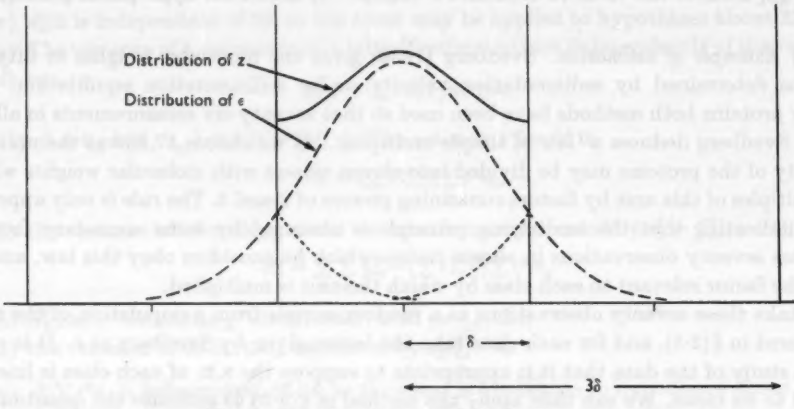


Fig. 2. Lumping the normal distrbution.

The *lumped variance* of the observations, s^2, is defined by

$$s^2 = \sum_{i=1}^{n} z_i^2/n.$$

We now consider the distribution of s^2/δ^2 on the quantum hypothesis when the subdistributions are normal. By the central limit theorem s^2/δ^2 is asymptotically normal, and we therefore derive its first two moments. The mean of s^2/δ^2 is the same as the mean of z^2/δ^2, and the variance of s^2/δ^2 is $1/n$ that of z^2/δ^2.

Now each z^2 is equal to ϵ^2 for $|\epsilon| \leqslant \delta$, and for all other ϵ these values are repeated with period 2δ. We may represent z^2 by a Fourier cosine series:

$$z^2 = \tfrac{1}{3}\delta^2 + \frac{4\delta^2}{\pi^2}\sum_{r=1}^{\infty}\frac{(-)^r}{r^2}\cos\left(\frac{r\pi\epsilon}{\delta}\right).$$

When calculating $E[z^2/\delta^2]$ we may integrate the series term-by-term, since it is a uniformly convergent series of continuous terms for all z. We obtain

$$E[z^2/\delta^2] = \frac{1}{3} + \frac{4}{\pi^2}\sum_{r=1}^{\infty}\frac{(-)^r}{r^2}\exp\left(-\frac{r^2\pi^2\sigma^2}{2\delta^2}\right).$$

Expanding z^4 as a Fourier series we obtain similarly

$$E[z^4/\delta^4] = \frac{1}{5} + \frac{8}{\pi^4}\sum_{r=1}^{\infty}\frac{(-)^r}{r^4}(r^2\pi^2 - 6)\exp\left(-\frac{r^2\pi^2\sigma^2}{2\delta^2}\right).$$

The two infinite series thus introduced are functions of $\frac{1}{2}\pi^2\sigma^2/\delta^2$ and have been tabulated by Newman (1934) in this form. They are integrals of one of the theta functions. We give in Table 1 some values of $E[z^2/\delta^2]$ and Var $[z^2/\delta^2]$ deduced from Newman's tables.

Now we write $E[z^2/\delta^2] = g(\sigma^2/\delta^2)$, the function g being tabulated below. It follows that s^2/δ^2 is an unbiased and consistent estimator of $g(\sigma^2/\delta^2)$:

$$g(\sigma^2/\delta^2) \approx s^2/\delta^2.$$

Since we require an estimator of σ^2 explicitly, we rewrite this equation,

$$\sigma^2 \approx \delta^2 g^{-1}(s^2/\delta^2) \equiv \delta^2 h(s^2/\delta^2).$$

This estimator is consistent and is recommended for general use rather than a maximum-likelihood estimator (which can be calculated if required) because of its simplicity. A table of $h(s^2/\delta^2)$ is needed when this estimator is used; it is given as Table 2, and has been obtained by interpolation in Table 1.

Table 1. *Mean and variance of z^2/δ^2 on the quantum hypothesis. The variance of s^2/δ^2 is $1/n$ that of z^2/δ^2*

$\frac{1}{2}\pi^2\sigma^2/\delta^2$	$E[z^2/\delta^2]$	Var $[z^2/\delta^2]$	$\frac{1}{2}\pi^2\sigma^2/\delta^2$	$E[z^2/\delta^2]$	Var $[z^2/\delta^2]$
0·0	0·0000	0·0000	1·0	0·1861	0·0516
·1	·0203	·0008	1·5	·2431	·0704
·2	·0405	·0033	2·0	·2785	·0795
·3	·0608	·0074	2·5	·3001	·0839
·4	·0809	·0129	3·0	·3132	·0861
0·5	0·1007	0·0194	3·5	0·3211	0·0873
·6	·1199	·0264	4·0	·3259	·0880
·7	·1382	·0334	4·5	·3288	·0883
·8	·1553	·0400	5·0	·3306	·0886
·9	·1713	·0461	∞	·3333	·0889

4. THE LUMPED VARIANCE TEST

(4·1) Suppose we are in the situation of § (3·2), so that the positions of the suspected modes have been given independently of the data. The observations are to be used to test whether there is real grouping about these modes. We neither know the weights p_i to be attached to each mode nor can we allot any observation to its correct subdistribution. The alternative to the quantum hypothesis has not been specified, but we wish to interpret the intuitive feeling that the alternative hypothesis implies no preference for the suspected modes and is perhaps rectangular or unimodal. We could rather not assume too much about the form of the subdistributions (e.g. normality), and a test which is in some sense distribution-free will therefore have advantages.

(4·2) In these circumstances a test using the lumped variance is appropriate. It has already been noted that neither the p_i nor the allocation of observations to the correct subdistributions are relevant to the distribution of s^2. Further, when the observations are clustered about the suspected modes s^2 is small in relation to δ^2. The tendency of the observations to be grouped at the modes is therefore measured by s^2/δ^2, and the form of the

subdistributions will be used only in this respect. It is, of course, true that if the true modes are not those suspected but are near them, and σ/δ is small, then s^2/δ^2 will again be small.

If y has a rectangular distribution, z has also a rectangular distribution between $-\delta$ and $+\delta$. If y has a smooth unimodal distribution whose spread is not small in comparison with δ, a little thought will show that z has approximately the rectangular distribution. Therefore, we take the hypothesis that z is rectangularly distributed between $-\delta$ and $+\delta$ to be our null hypothesis. On this rectangular hypothesis s^2/δ^2 is the mean of n independent variates each of which can be shown to have mean $\frac{1}{3}$ and variance $\frac{4}{45}$. These are also the limits of the mean and variance of z^2/δ^2 on the quantum hypothesis as $\sigma/\delta \to \infty$. By the central limit theorem, s^2/δ^2 is approximately normally distributed with mean $\frac{1}{3}$ and variance $\frac{4}{45}n$ when n is large; for $n \geqslant 20$ the approximation may be expected to be good.

Table 2. $h(s^2/\delta^2)$. *An estimate of* σ^2 *is given by* $\delta^2 h(s^2/\delta^2)$

s^2/δ^2	$h(s^2/\delta^2)$	s^2/δ^2	$h(s^2/\delta^2)$
0	0·000	0·16	0·167
0·01	·010	·17	·180
·02	·020	·18	·194
·03	·030	·19	·208
·04	·040	·20	·223
0·05	0·050	0·21	0·238
·06	·060	·22	·255
·07	·070	·23	·274
·08	·080	·24	·295
·09	·090	·25	·318
0·10	0·100	0·26	0·344
·11	·111	·27	·374
·12	·122	·28	·409
·13	·132	·29	·452
·14	·143	·30	·506
·15	·155	·31	·577
		·32	·690

On the quantum hypothesis s^2/δ^2 has expectation less than $\frac{1}{3}$. As σ^2 increases the distribution, although grouped, becomes indistinguishable from one of the class of alternatives.

(4·3) The lumped variance test is a one-sided test of the rectangular hypothesis, and gives the probability that the value of s^2/δ^2 found, or a lower value, would occur by chance if z were rectangularly distributed. It will be noted that it is the alternative, and not the quantum hypothesis which is tested.

The significance points, for probability levels 0·05, 0·01 and 0·001 and $n = 20\,(5)\,100\,(50)\,1000$ are given in Table 3. If the value of s^2/δ^2 found is less than the appropriate value in this table, the rectangular hypothesis is rejected and the hypothesis of grouping, or quantum hypothesis, is accepted. It will be remembered that the rectangular hypothesis may be accepted if the variance σ^2 of the subdistributions is large in relation to δ^2.

(4·4) *Example: distribution of length in* Labidocera euchaeta. Seymour Sewell (1912) measured the total lengths of large numbers of Copepoda with a view to testing whether

they followed Brook's law (Fowler, 1909) as stated for Somatopoda and Ostracoda. From his data we extract Table 4, of the distribution of length of females of the species *Labidocera euchaeta*. Sample A (Seymour Sewell's fig. 1) consists of 497 specimens from the Rangoon

Table 3. *Lumped variance test. If s^2/δ^2, calculated from n observations, is less than the value tabulated at probability level P, the quantum hypothesis is accepted at that probability level*

n	$P=0{\cdot}05$	$P=0{\cdot}01$	$P=0{\cdot}001$
20	0·2237	0·1782	0·1273
25	·2353	·1946	·1490
30	·2438	·2067	·1651
35	·2504	·2161	·1776
40	·2558	·2237	·1878
45	0·2602	0·2299	0·1960
50	·2640	·2352	·2030
55	·2672	·2398	·2091
60	·2700	·2438	·2144
65	·2725	·2473	·2191
70	0·2747	0·2504	0·2232
75	·2767	·2532	·2269
80	·2785	·2558	·2303
85	·2801	·2581	·2334
90	·2816	·2602	·2362
95	·2830	·2622	·2388
100	0·2843	0·2640	0·2412
150	·2933	·2767	·2581
200	·2987	·2843	·2682
250	·3023	·2895	·2751
300	·3050	·2933	·2801
350	0·3071	0·2963	0·2841
400	·3088	·2987	·2873
450	·3102	·3006	·2899
500	·3114	·3023	·2921
550	·3124	·3038	·2940
600	0·3133	0·3050	0·2957
650	·3141	·3061	·2972
700	·3148	·3071	·2985
750	·3154	·3080	·2997
800	·3160	·3088	·3008
850	0·3165	0·3095	0·3017
900	·3170	·3102	·3026
950	·3174	·3108	·3034
1000	·3178	·3114	·3042

River Estuary, and sample B (his fig. 2) of 157 collected after sample A from Chittagong. We shall use these data to answer the question 'Do the specimens from Chittagong show any tendency to occur relatively more frequently at the modes from the first sample?'

It is expected from previous work that when length is taken on a logarithmic scale the data will be grouped about approximately equally spaced modes (any change in the growth factor is here being neglected). Because the groups overlap and no other information is available, the positions of the population modes of sample A cannot be estimated by the methods of § (2). When the experimental modes at 15, 22, 30, 42, 54 and 65 units of length (the modes at 20, 36 and 49 units are seen from a figure to be spurious) are plotted on a

Table 4. *Distribution of lengths of females*, Labidocera euchaeta (*unit of length* 0·04 mm.). *Semour Sewell's data*

Length	Sample A	Sample B	Length	Sample A	Sample B
13	2	0	43	15	6
14	3	0	44	5	3
15	3	0	45	6	3
16	3	0	46	9	2
17	1	0	47	16	2
18	0	0	48	18	2
19	1	0	49	23	2
20	3	0	50	21	8
21	2	5	51	25	11
22	4	0	52	25	12
23	1	0	53	28	10
24	2	0	54	31	9
25	4	0	55	19	9
26	5	1	56	9	7
27	10	2	57	4	1
28	10	5	58	2	2
29	12	4	59	1	0
30	14	2	60	0	1
31	12	0	61	1	0
32	8	0	62	0	1
33	5	0	63	1	0
34	1	1	64	7	2
35	2	1	65	12	2
36	6	3	66	9	0
37	5	4	67	7	1
38	7	7	68	4	0
39	12	6	69	3	0
40	18	5	70	1	1
41	18	6	71	1	1
42	20	7			
			Totals	497	157

logarithmic scale against $r = 0, 1, \ldots, 5$ they lie approximately on a straight line. Fitting a line by eye we conclude that sample A suggests modes at 1·218 (0·126) 1·848 on a logarithmic scale.

We next require to test whether sample B supports this grouping, and do so by the lumped variance test. Taking logarithms, we calculate the square of the distance of each observation of sample B from the nearest mode deduced from sample A. The sum of these squared distances, divided by $n\delta^2 = 157\,(0{\cdot}126/2)^2$, is 0·177. Comparison of this value with Table 3 shows that departure from the value expected on the rectangular hypothesis is highly

significant. We conclude that sample B supports strongly the grouping we have deduced from sample A, i.e. that grouping is real and occurs near the modes given above.

If an estimate of the scatter about these modes is required, and the assumptions of § (3·2) are justified, the method of that section gives an estimate of the variance about each mode on a logarithmic scale. Since s^2/δ^2 is 0·177, and δ^2 is $(0{\cdot}063)^2$, the variance is estimated as $\delta^2 h(s^2/\delta^2) = (0{\cdot}063)^2(0{\cdot}189) = 0{\cdot}000750$.

(4·5) *Example: choosing a random point.* The following simple experiment is described because it exemplifies the use of lumped variance. It was designed to test whether the points on a line chosen 'at random' by a number of people were in fact grouped about certain regular points on the line.

Twenty subjects were chosen from workers at B.C.U.R.A. and each was presented with a sheet of paper on which were drawn eight parallel lines of length 12 cm. Each subject was

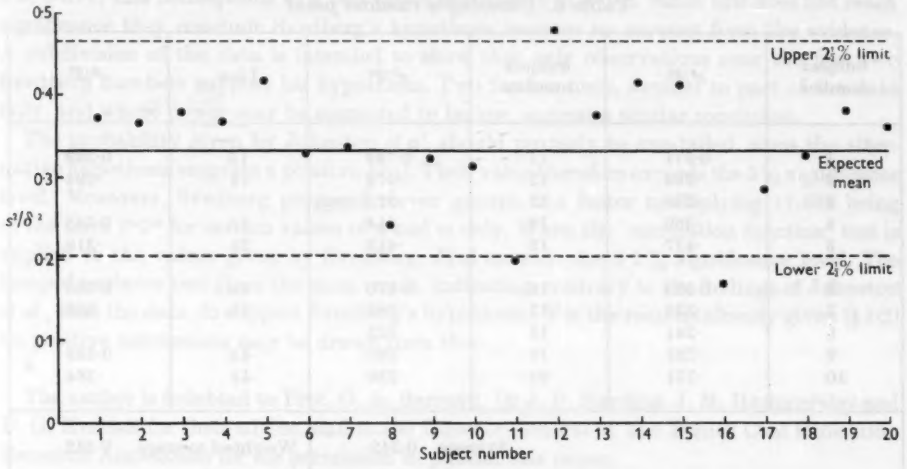


Fig. 3. Choosing a random point. Control chart of s^2/δ^2.

then asked to mark points on the lines at random, i.e. without systematic placing. He was to make one mark on each of the first two lines, two, three and four on each of the successive pairs of lines, making twenty marks in all. Alternate subjects marked their lines in the reverse order.

The points about which systematic grouping was expected were those which divide the line into $(n+1)$ equal segments when n marks were made. Thus when one mark was made it was expected the end-points and centre would be preferred, when two marks were made the end-points and thirds, and so on. In this situation the lumped variance test is applicable. When n points are marked and l is the length of the line, δ is $l/2(n+1)$, and z is the distance of a mark from the nearest 'preferred' point as defined above.

The values for s^2/δ^2 found in the experiment are given in Table 5, and are shown also for subjects in the form of a control chart. In Table 5, 1*a* denotes the first line on which one mark was made, 1*b* the second line on which one was made, and so on.

The conclusion drawn is that no evidence was found that the subjects did not choose points on the lines at random; the data are consistent with the hypothesis that all points

on the line are equally likely to be chosen. Clearly more extensive experiments and other tests may reverse this conclusion. One subject has s^2/δ^2 at about the 99% point, i.e. his marks tend to avoid the regular points, and two subjects have s^2/δ^2 below the 3% point, i.e. their marks do cluster at these points, but these are the extremes in a sample of twenty. None of the s^2/δ^2 for lines, nor the general average, is significantly different from $\frac{1}{3}$ as expected on the rectangular hypothesis. Further, the variance of s^2/δ^2 for subjects is 0·0056; this does not differ significantly from the variance expected on the rectangular hypothesis, 0·0044.

(4·6) Finally, we consider the statistical tests applied by Johnston *et al.* (1945) to Svedberg's (1939) data. We have already noted, in § (1·2), the difficulty in testing a hypothesis which is suggested by the data used in the test. Suppose, however, that a lumped variance

Table 5. *Choosing a random point*

Subject number	s^2/δ^2	Subject number	s^2/δ^2	Line	s^2/δ^2
1	0·371	11	0·199	1*a*	0·369
2	·364	12	·479	1*b*	·294
3	·379	13	·373		
4	·366	14	·414	2*a*	0·345
5	·417	15	·413	2*b*	·316
6	0·328	16	0·170	3*a*	0·360
7	·336	17	·286	3*b*	·309
8	·241	18	·325		
9	·321	19	·380	4*a*	0·348
10	·311	20	·359	4*b*	·364
		Average	0·342	Weighted average	0·342

test is applied to such data, using the modes suggested by or calculated from the data, and that the value of s^2/δ^2 found is not significantly less than the value $\frac{1}{3}$ expected on the rectangular hypothesis. Then we would correctly reject the quantum hypothesis and accept the class of distribution defined in § (4·2). The reason for doing so is that when the modes are suggested by the data the value of s^2/δ^2 will therefore be small, corresponding to the grouping of the data about these modes. In such circumstances it would be wrong to accept the quantum hypothesis solely because s^2/δ^2 is less than a conventional significance point, whereas it is proper to reject the quantum hypothesis if s^2/δ^2 does not reach significance.

The 'correlation function' test applied by Johnston *et al.* is similar to the lumped variance test proposed in § (4·3); their statistic stands in much the same relation to the lumped variance as the mean deviation does to the variance. The lumped variance has the advantage that it is useful in the estimation of variance by the method of § (3·2). Their statistic is, in the notation defined above,

$$F(n) = 1 - 2 \sum_{i=1}^{n} \{|z_i|/(n\delta)\}.$$

It is distributed approximately normally about zero with variance $1/(3n)$ on the rectangular hypothesis, and a positive value indicates grouping about the supposed modes.

They extend this definition of $F(n)$ to those cases in which the modes are not equally spaced, taking as the δ corresponding to each z_i half the distance between the modes on either side of z_i. A similar extension is, of course, possible for the statistic $\sum_{i=1}^{n} z_i^2/(n\delta^2)$ of § (3·2). The rectangular hypothesis they specify in this case to be a rectangular distribution of y between each mode; this is unnecessarily restricted, for we can consider the class of all distributions of y which correspond to a rectangular distribution of z.

Johnston *et al.* apply their test to the fifty-six values given by Svedberg, averaging the results of two determinations when these are available. They test for grouping about the values $2^n3^m \times 17{,}600$ for all integral n and m, and obtain for $F(n)$ the value 0·137 with S.D. 0·077; this corresponds to a two-tailed probability of 0·076. Since this does not reach significance they conclude Svedberg's hypothesis receives no support from the evidence. A subdivision of the data is intended to show that only observations near the first two Svedberg numbers support his hypothesis. Two further tests, applied to part of the data only, and whose power may be suspected to be low, suggest a similar conclusion.

The probability given by Johnston *et al.* should properly be one-tailed, since the alternative hypothesis suggests a positive $F(n)$. Their value therefore exceeds the 5% significance level. Moreover, Svedberg proposed eleven groups, the factor multiplying 17,600 being of the form 2^n3^m for certain values of n and m only. When the 'correlation function' test is applied to the values given by Svedberg, $F(n)$ exceeds the 0·1% significance level. The lumped variance test gives the same result, indicating, contrary to the findings of Johnston *et al.*, that the data do support Svedberg's hypothesis. For the reasons already given (§ 1·2) no positive conclusions may be drawn from this.

The author is indebted to Prof. G. A. Barnard, Dr J. P. Harding, J. M. Hammersley and D. G. Kendall for their advice, and to the Director-General of The British Coal Utilisation Research Association for his permission to publish this paper.

REFERENCES

BROOKS, W. K. (1886). *Report on the Scientific Results of the Exploring Voyage of H.M.S. Challenger*, 1873–76, **16**, (Zoological Section), *Stomatopoda*, 1.
FOWLER, G. H. (1909). *Trans. Linn. Soc.* (*Zool.*), 2nd ser. pt. 9.
GRANT, P. J. (1952). *Proc. Phys. Soc., Lond.*, A, **65**, 150.
HAMMERSLEY, J. M. (1950). *J. R. Statist. Soc.* B, **12**, 192.
HAMMERSLEY, J. M. & MORTON, K. W. (1954). *J. R. Statist. Soc.* B, **16**, 23.
JOHNSTON, J. P., LONGUET-HIGGINS, H. C. & OGSTON, A. G. (1945). *Trans. Faraday Soc.* **41**, 588.
NEWMAN, A. B. (1934). *Trans. Amer. Inst. Chem. Engrs*, **30**, 598.
PEARSON, E. S. & CHANDRASEKHAR, C. (1936). *Biometrika*, **28**, 308.
PEARSON, K. (1894). *Phil. Trans.* A, **185**, 71.
PRZIBRAM, H. & MEGUŠAR, F. (1912). *Roux Arch. EntwMech. Organ.* **34**, 680.
SEYLER, C. A. (1943). *J. Inst. Fuel*, **16**, 134.
SEYMOUR SEWELL, R. B. (1912). *Rec. Indian Mus.* **7**, 313.
SVEDBERG, T. (1939). *Proc. Roy. Soc.* A, **170**, 40 and B, **127**, 1.
THOM, A. (1955). *J. R. Statist. Soc.* B. (in the Press).
WIGGLESWORTH, V. D. (1942). *Principles of Insect Physiology*, 2nd ed. London: Methuen.

THE TRUNCATED NEGATIVE BINOMIAL DISTRIBUTION

By M. R. SAMPFORD

Agricultural Research Council Unit of Statistics, University of Aberdeen

1. Introduction

The negative binomial distribution has been discussed by, *inter alia*, Greenwood & Yule (1920), Fisher (1941), Haldane (1941), Anscombe (1950) and Bliss & Fisher (1953), and is extensively used for the description of data too heterogeneous to be fitted by a Poisson distribution. Observed samples, however, may be truncated, in the sense that the number of individuals falling into the zero class cannot be determined. For example, if chromosome breaks in irradiated tissue can occur only in those cells which are at a particular stage of the mitotic cycle at the time of irradiation, a cell can be demonstrated to have been at that stage only if breaks actually occur. Thus in the distribution of breaks per cell, cells not susceptible to breakage are indistinguishable from susceptible cells in which no breaks occur.

Methods for estimation of the parameters of the truncated distribution are considered in this paper. The corresponding problem of estimation of the truncated Poisson distribution has been discussed by David & Johnson (1952), who also discuss the present problem.

2. The moments of the truncated distribution

The negative binomial distribution has the form (in Fisher's notation)

$$P(r) = \frac{(k+r-1)!}{(k-1)!\,r!}\,\frac{p^r}{(1+p)^{k+r}} \quad (r = 0, 1, \ldots;\ p, k > 0), \tag{1}$$

so that

$$P(0) = 1/(1+p)^k.$$

To obtain the corresponding probabilities for the truncated distribution the form (1) must be divided by $(1-P(0))$; writing

$$\varpi = 1/(1+p), \quad \eta = 1-\varpi,$$

it follows that

$$P_t(r) = \frac{\varpi^k}{1-\varpi^k}\,\frac{(k+r-1)!}{(k-1)!\,r!}\,\eta^r \quad (r = 1, 2, \ldots). \tag{2}$$

The factorial moments of this distribution are

$$\mu'_{(j)} = \frac{(k+j-1)!\,\eta^j}{(k-1)!\,\varpi^j(1-\varpi^k)}, \tag{3}$$

whence

$$\mu'_1 = \frac{k\eta}{\varpi(1-\varpi^k)},$$

$$\mu'_2 = \frac{k\eta + k^2\eta^2}{\varpi^2(1-\varpi^k)},$$

$$\mu'_3 = \frac{k\eta + k(3k+1)\,\eta^2 + k^3\eta^3}{\varpi^3(1-\varpi^k)},$$

$$\mu'_4 = \frac{k\eta(6-6\varpi+\varpi^2) + k^2\eta^2(11-4\varpi) + 6k^3\eta^3 + k^4\eta^4}{\varpi^4(1-\varpi^k)}.$$

3. ESTIMATION BY MOMENTS

David & Johnson (1952) suggest that the use of estimates of less than maximum efficiency is justifiable only if they are directly obtainable as explicit solutions of easily constructed equations. In discussing the truncated negative binomial, therefore, they do not consider estimates based on the first two sample moments, which do not provide explicit solutions, but confine their attention to a method using certain ratios of the first three moments, i.e. sample estimates of μ_2'/μ_1' and μ_3'/μ_1'. The estimates (only that of p is discussed in detail) are obtained easily enough, but, in consequence of the introduction of the third sample moment, are extremely inefficient. (For example, the efficiency for values of the parameters equivalent to $k = 1$, $\varpi = 0{\cdot}5$ is as low as $1{\cdot}7\,\%$.) David & Johnson therefore abandon completely the use of moments and recommend the maximum-likelihood method for use in all cases.

Whether, in fact, any particular inefficient procedure is acceptable can only depend on the loss of information resulting from its use, the time saved by it, and the relative costs of the time or labour spent on observation and on analysis. Thus if an experiment involves observations on hundreds of experimental animals, made over a period of several years, ten or even a hundred hours of calculation may be dearly saved at the cost of $10\,\%$ of the information so laboriously accumulated. If, on the other hand, the observations are made easily and at no great cost, the use of a convenient but statistically 'inefficient' method of analysis, coupled with an appropriate increase in sample size, may be far more 'efficient' than a tedious maximum-likelihood calculation, in the sense of giving the same amount of information at a lower cost. In this section I give a trial-and-error method for solving the moment equations using the first two moments. This method, though not explicit, does not take many minutes to carry through. It certainly entails far less labour than the maximum-likelihood calculations, and the estimates obtained have a percentage efficiency of 80 or upwards for all but the most unfavourable combinations of parameters: for $k = 1$, $\varpi = 0{\cdot}5$ the efficiency is $77{\cdot}5\,\%$ (§ 6, Table 2), as compared with the $1{\cdot}7\,\%$ of the 'three-moment' estimate. This method can thus be recommended as a reasonable alternative to the maximum-likelihood calculations, in circumstances where a method of less than $100\,\%$ efficiency seems likely to prove acceptable.

Equating the population mean and variance to the corresponding sample values gives equations for k and ϖ:

$$\left.\begin{aligned} \frac{k\eta}{\varpi(1-\varpi^k)} &= m, \\ \frac{k\eta(1+k\eta)}{\varpi^2(1-\varpi^k)} &= m^2+s^2. \end{aligned}\right\} \tag{4}$$

These equations can be solved by trial and error, for which purpose they are most conveniently expressed in the form

$$\left.\begin{aligned} k &= \left(m+\frac{s^2}{m}-1\right)\frac{\varpi}{1-\varpi}-1, \\ \xi &= m\varpi^{k+1}+\frac{s^2}{m}\varpi-1 = 0. \end{aligned}\right\} \tag{5}$$

The value of ξ can be evaluated for selected values of ϖ, and a solution, if one exists, reached by successive linear interpolations.

To investigate whether a solution exists, and to simplify still further the computations required, we consider the function

$$\phi(x) = \frac{-x \log_e x}{1-x}. \tag{6}$$

This function is tabulated, for values of x between 0 and 1, in Table 4, and can be shown to possess the properties

$$\phi(x) \to 0, 1 \quad \text{as} \quad x \to 0, 1,$$
$$\phi'(x) > 0, \quad \phi''(x) < 0, \quad 0 < x < 1.$$

The second of equations (5) then takes the form

$$\zeta(\varpi) = m \exp\left\{-\left(m + \frac{s^2}{m} - 1\right)\phi(\varpi)\right\} + \frac{s^2}{m}\varpi = 1. \tag{7}$$

If all the observed values equal 1, $m = 1$ and $s^2 = 0$, and $\zeta \equiv 1$. Thus in this situation the moment estimation procedure, naturally enough, fails. In all further discussion, therefore, m will be assumed > 1.

It is easily verified that

$$\varpi_0 = 1 \Big/ \left(m + \frac{s^2}{m}\right) \quad (< 1)$$

satisfies equation (7). However, the corresponding value of k, obtained from the first of equations (5), is 0. ϖ_0, therefore, and any lower values of ϖ (corresponding to negative values of k) are inadmissible: we require a solution of (7) in the range

$$\varpi_0 < \varpi < 1.$$

We have $\zeta(0) = m > 1, \quad \zeta(\varpi_0) = 1, \quad \zeta'' > 0 \quad (0 < \varpi < 1).$

(The result on ζ'' follows from that on ϕ'' quoted above.) There exists, therefore, at most one solution of (7) other than ϖ_0 in the range $0 < \varpi < 1$. The condition that there shall be another solution less than 1 is

$$\zeta(1) > 1,$$

the condition that it shall be greater than ϖ_0 is

$$\zeta'(\varpi_0) < 0.$$

These inequalities reduce to limitations on the relative magnitudes of m and s^2, most conveniently expressed in the form

$$1 - \exp\left\{-\left(m + \frac{s^2}{m} - 1\right)\right\} < \left(m + \frac{s^2}{m} - 1\right) \Big/ m < \log\left(m + \frac{s^2}{m}\right). \tag{8}$$

Neither of these inequalities is necessarily satisfied; it is quite possible to construct samples in which they are not. In particular, the lower inequality is not satisfied if $s^2 = 0$, whatever the value of m (reasonably enough; we should hardly expect to get sensible estimates by equating an essentially positive function to zero). However, provided very small samples are avoided, and provided m is not too near 1, (8) will probably be satisfied.

Equation (7) can be solved either iteratively, writing

$$\varpi_{i+1} = \frac{m}{s^2}\left\{1 - m\exp\left[-\left(m + \frac{s^2}{m} - 1\right)\phi(\varpi_i)\right]\right\}, \tag{9}$$

or by trial and error. Inasmuch as the solution must lie on that part of the ζ curve for which $\zeta' > 0$, between the (unknown) minimum of ζ and the point $\varpi = 1$, it is most convenient to take as the first trial value of ϖ either 1, or, if it is less than 1, $\varpi = m/s^2$ (clearly $\zeta > 1$ for this value), and to work down to the solution. This approach will ensure the convergence of the iterative procedure (9), and should minimize the amount of labour required for the trial-and-error procedure. In the latter case time will probably be saved by calculating, for the first trial value of ϖ, the slope

$$\zeta'(\varpi) = \frac{s^2}{m} - m\left(m + \frac{s^2}{m} - 1\right)\left[\frac{\phi(\varpi) - \varpi}{\varpi(1-\varpi)}\right]\exp\left\{-\left(m + \frac{s^2}{m} - 1\right)\phi(\varpi)\right\}. \tag{10}$$

Except for the function in square brackets ($= \phi'(\varpi)$) all the terms used in calculating this expression will already have been evaluated in the calculation of ζ. This slope can be used in choosing the second trial value, after which one or two linear interpolations or extrapolations should be sufficient. If $\varpi = 1$ is taken as the initial approximate, $\phi'(\varpi)$ should be replaced by its limiting value of $\frac{1}{2}$.

4. Example of estimation by moments

In an investigation into chromosome breakage, the following sample distribution of breaks per cell was obtained:

$$r = 1(11),\ 2(6),\ 3(4),\ 4(5),\ 6(1),\ 8(2),\ 9,\ 11,\ 13.$$

$$n = 32, \quad \Sigma r = 110, \quad \Sigma r^2 = 686.$$

$$m = 3{\cdot}4375, \quad s^2 = 9{\cdot}9315, \quad s^2/m = 2{\cdot}8892.$$

Equation (7) is thus

$$3{\cdot}4375\, e^{-5{\cdot}3267\phi} + 2{\cdot}8892\varpi = 1.$$

Taking as the first trial value

$$\varpi = 1/2{\cdot}8892 = 0{\cdot}3461,$$

we obtain from Table 4

$$\phi = 0{\cdot}5616,$$

whence

$$(\phi - \varpi)/\varpi(1-\varpi) = 0{\cdot}9522.$$

Then

$$5{\cdot}3267\phi = 2{\cdot}9915,$$

and from tables of the negative exponential function, or of natural logarithms,

$$e^{-2{\cdot}9915} = 0{\cdot}05021, \quad 3{\cdot}4375\, e^{-2{\cdot}9915} = 0{\cdot}1726,$$

whence

$$\zeta = 0{\cdot}1726 + 1{\cdot}0000 = 1{\cdot}1726$$

and

$$\zeta' = 2{\cdot}8892 - (0{\cdot}1726 \times 5{\cdot}3267 \times 0{\cdot}9522) = 2{\cdot}0138,$$

suggesting, as the next trial value,

$$\varpi = 0{\cdot}3461 - 0{\cdot}1726/2{\cdot}0138 = 0{\cdot}2604.$$

The remainder of the trial-and-error solution, shown in Table 1 (*a*), leads to the value

$$\varpi = 0{\cdot}2346,$$

whence

$$k = 5{\cdot}3267 \times 0{\cdot}2346/0{\cdot}7654 - 1 = 0{\cdot}633.$$

Alternatively, by the iterative method, starting from the same initial value, the second approximation

$$\varpi_2 = (1 - 0{\cdot}1726)/2{\cdot}8892 = 0{\cdot}2864.$$

The remainder of the calculations, leading after sixteen cycles to the same solution, are shown in Table 1 (*b*).

Table 1. *Solution of the moment estimation equation for the example of* § 4

(*a*) By trial and error

ϖ	ϕ	$3{\cdot}4375\,e^{-5{\cdot}3267\phi}$	ζ	ζ'
0·3461	0·5616	0·1726	1·1726	2·0137
0·2604	0·4737	0·2757	1·0280	
0·2438	0·4550	0·3046	1·0090	
0·2359	0·4459	0·3197	1·0013	
0·2346	0·4444	0·3222	1·0000	

(*b*) By iteration

ϖ	ϕ	$3{\cdot}4375\,e^{-5{\cdot}3267\phi}$
0·3461	0·5616	0·1726
0·2864	0·5018	0·2374
0·2639	0·4776	0·2700
0·2527	0·4651	0·2886
0·2462	0·4578	0·3000
0·2423	0·4533	0·3073
0·2397	0·4503	0·3123
0·2380	0·4484	0·3154
0·2370	0·4472	0·3175
0·2362	0·4463	0·3190
0·2357	0·4457	0·3200
0·2353	0·4452	0·3209
0·2350	0·4448	0·3216
0·2348	0·4446	0·3219
0·2347	0·4445	0·3221
0·2346	0·4444	0·3222
0·2346		

(Intermediate stages of the calculations are shown in Table 1 for the sake of the example. In practice either calculation can be carried through on a desk computer, recording only the successive values of ϖ and, for the trial-and-error process, ζ.)

Four decimal places have been retained in the estimation of ϖ as a demonstration of the degree of precision to which the moment equations can be solved by this method. In fact, for this example the variances (§ 5) are so large that there is no practical point in retaining more than two decimal places for ϖ and one for k.

5. The variances of the moment estimates

The asymptotic variances and covariances of the moment estimates of ϖ and k are given by the formulae

$$J^2V(\varpi) = \left(\frac{\partial\mu_2'}{\partial k}\right)^2 V(m_1') - 2\frac{\partial\mu_1'}{\partial k}\frac{\partial\mu_2'}{\partial k}\operatorname{cov}(m_1', m_2') + \left(\frac{\partial\mu_1'}{\partial k}\right)^2 V(m_2'),$$

$$J^2\operatorname{cov}(\varpi, k) = -\frac{\partial\mu_2'}{\partial k}\frac{\partial\mu_2'}{\partial\varpi}V(m_1') + \left(\frac{\partial\mu_1'}{\partial k}\frac{\partial\mu_2'}{\partial\varpi} + \frac{\partial\mu_1'}{\partial\varpi}\frac{\partial\mu_2'}{\partial k}\right)\operatorname{cov}(m_1', m_2') - \frac{\partial\mu_1'}{\partial k}\frac{\partial\mu_1'}{\partial\varpi}V(m_2'),$$

$$J^2V(k) = \left(\frac{\partial\mu_2'}{\partial\varpi}\right)^2 V(m_1') - 2\frac{\partial\mu_1'}{\partial\varpi}\frac{\partial\mu_2'}{\partial\varpi}\operatorname{cov}(m_1', m_2') + \left(\frac{\partial\mu_1'}{\partial\varpi}\right)^2 V(m_2'),$$

where

$$\frac{\partial\mu_1'}{\partial k} = \frac{\mu_1'}{k}\{1-\phi(\varpi^k)\},$$

$$\frac{\partial\mu_1'}{\partial\varpi} = -\frac{\mu_1'}{\varpi\eta}\{1-\varpi^{k+1}\mu_1'\},$$

$$\frac{\partial\mu_2'}{\partial k} = -\frac{\mu_1'}{k\varpi} + \frac{\mu_2'}{k}\{2-\phi(\varpi^k)\},$$

$$\frac{\partial\mu_2'}{\partial\varpi} = \frac{\mu_1'}{\varpi\eta} - \frac{\mu_2'}{\varpi\eta}\{2-\varpi^{k+1}\mu_1'\},$$

$$J = \frac{\partial\mu_1'}{\partial k}\frac{\partial\mu_2'}{\partial\varpi} - \frac{\partial\mu_1'}{\partial\varpi}\frac{\partial\mu_2'}{\partial k},$$

and $\quad nV(m_1') = (\mu_2' - \mu_1'^2), \quad n\operatorname{cov}(m_1', m_2') = (\mu_3' - \mu_1'\mu_2'), \quad nV(m_2') = (\mu_4' - \mu_2'^2).$

Inserting into these formulae the moment estimates obtained in the example of the previous section,

$$\varpi = 0{\cdot}2346, \quad k = 0{\cdot}633,$$

we have ($\log_e\varpi = -1{\cdot}44988$, $\varpi^k = 0{\cdot}3994$):

$$\mu_1' = 3{\cdot}438578, \quad \mu_2' = 21{\cdot}758578, \quad \mu_3' = 215{\cdot}773952, \quad \mu_4' = 2941{\cdot}290671,$$

whence $\quad V(m_1') = 0{\cdot}31046123, \quad \operatorname{cov}(m_1', m_2') = 4{\cdot}4048558, \quad V(m_2') = 77{\cdot}120467.$

Also $\quad \partial\mu_1'/\partial k = 2{\cdot}1164, \quad \partial\mu_1'/\partial\varpi = -12{\cdot}9798, \quad \partial\mu_2'/\partial k = 24{\cdot}6106, \quad \partial\mu_2'/\partial\varpi = -184{\cdot}1591$

$$J = -70{\cdot}31365336,$$

whence $\quad V(\varpi) = 0{\cdot}015091, \quad \operatorname{cov}(\varpi, k) = 0{\cdot}08125, \quad V(k) = 0{\cdot}4983.$

6. The efficiency of the moment procedure

It has been shown by Fisher (1941) that the method of moments may be seriously inefficient for the estimation, from complete data, of the parameters of the negative binomial distribution. Investigation of the corresponding efficiency for estimation from truncated data therefore seems desirable.

The determinant of the variance-covariance matrix of the moment estimates reduces to

$$\frac{2(k+1)\varpi^2(1-\varpi^k)^3}{n^2\eta} \times \frac{\{1-\varpi^k[1+k\eta+\frac{1}{2}k(k+1)\eta^2]\}}{\{1-\varpi^k[1-k^2\eta-k(k+1)\log\varpi]\}^2}.$$

The determinant of the information matrix for maximum efficiency is

$$\frac{n^2\{1-\varpi^k[1+k\eta]\}}{\varpi^2\eta(1-\varpi^k)^3}\left\{\sum_{r=2}^{\infty}\frac{\eta^r}{r}\frac{(r-1)!\,k!}{(k+r-1)!} - \frac{k\varpi^k(\eta+\log\varpi)^2}{\{1-\varpi^k[1+k\eta]\}}\right\},$$

whence the reciprocal of the efficiency of the moment method, given by the product of these two expressions, is

$$\frac{1}{E} = \frac{\{1-\varpi^k[1+k\eta+\frac{1}{2}k(k+1)\eta^2]\}\{1-\varpi^k[1+k\eta]\}}{\{1-\varpi^k[1-k^2\eta-k(k+1)\log\varpi]\}^2} \times\left\{2\sum_{r=2}^{\infty}\frac{\eta^{r-2}}{r}\frac{(r-1)!\,(k+1)!}{(k+r-1)!}-\frac{2k(k+1)\varpi^k}{[1-\varpi^k(1+k\eta)]}\left[1+\frac{\log\varpi}{\eta}\right]^2\right\}.$$

Table 2 shows the percentage efficiency of the moment method for selected values of k, and of the mean of the complete distribution. From this it is clear that even for quite small means there may be a serious loss of efficiency for low values of k.

For the example of § 4, k was estimated as 0·633, and the mean of the complete distribution $= k\eta/\varpi = 2{\cdot}065$, from which estimation by moments would appear to be about 70 % efficient in this case.

Table 2. *Percentage efficiency of the moment method of estimation, for selected values of k and of the mean of the complete distribution*

Mean \ k	0·5	1	2	3	4	5
0·5	82·4	90·6	95·7	97·5	98·3	98·8
1·0	74·6	84·7	92·4	95·4	96·9	97·8
2·0	66·1	77·5	87·5	92·0	94·4	95·9
5·0	55·9	67·3	79·2	85·5	89·2	91·6

7. Estimation by maximum likelihood

In view of the results of the previous section, it seems desirable to consider the maximum-likelihood estimation procedure in some detail.

The log likelihood is

$$\log L = \sum_1^{\infty} n_r \log\left\{\frac{\varpi^k}{1-\varpi^k}\frac{(k+r-1)!}{(k-1)!\,r!}(1-\varpi)^r\right\}$$

$$= nk\log\varpi - n\log(1-\varpi^k) + \sum_1^{\infty} rn_r\log(1-\varpi) - \sum_1^{\infty} n_r\log r! + \sum_{r=1}^{\infty} n_r \sum_{j=1}^{r}\log(k+j-1),$$

giving the maximum-likelihood equations

$$\frac{nk}{\varpi(1-\varpi^k)} - \frac{nm}{(1-\varpi)} = 0 \tag{11}$$

and

$$\frac{n\log\varpi}{1-\varpi^k} + \sum_{r=1}^{\infty} n_r \sum_{j=1}^{r}\frac{1}{(k+j-1)} = 0, \tag{12}$$

which, following Haldane (1941), is conveniently rewritten as

$$\frac{n\log\varpi}{1-\varpi^k} + \sum_{j=1}^{R}(k+j-1)^{-1}\sum_{i=j}^{R} n_i = 0, \tag{13}$$

where R is the highest observed value of r.

These equations are easily soluble by the usual maximum-likelihood iterative procedure. The components of the information matrix are the expected values of the quantities

$$\left.\begin{aligned} -\frac{\partial^2 \log L}{\partial \varpi^2} &= \frac{nk[1-(k+1)\varpi^k]}{\varpi^2(1-\varpi^k)^2} + \frac{nm}{(1-\varpi)^2}, \\ -\frac{\partial^2 \log L}{\partial \varpi \partial k} &= -\frac{n[1-(1-k\log \varpi)\varpi^k]}{\varpi(1-\varpi^k)^2}, \\ -\frac{\partial^2 \log L}{\partial k^2} &= \sum_{j=1}^{R}(k+j-1)^{-2}\sum_{i=j}^{R} n_i - \frac{n(\log \varpi)^2 \varpi^k}{(1-\varpi^k)^2}. \end{aligned}\right\} \tag{14}$$

The iteration, however, is most conveniently carried out using the quantities (14) themselves, rather than their expected values. An example of the calculations for this method is given at the end of the next section.

8. An alternative method for the solution of the maximum-likelihood equations

From equations (11) and (13) it follows that

$$\phi(\hat{\varpi}) = -\frac{\hat{\varpi}\log_e \hat{\varpi}}{1-\hat{\varpi}} = \frac{\hat{k}}{nm}\sum_{j=1}^{R}(\hat{k}+j-1)^{-1}\sum_{r=j}^{R} n_r, \tag{15}$$

where $\hat{\varpi}$ and $\hat{k}$ are maximum-likelihood estimates, and m is the sample mean. Equation (11) can be rewritten

$$\psi(\hat{\varpi}, \hat{k}) = \frac{\hat{\varpi}(1-\hat{\varpi}^k)m}{\hat{k}(1-\hat{\varpi})} = 1. \tag{16}$$

For a given value of k, equation (15) can be used to evaluate the corresponding ϕ and hence, from Table 4, ϖ; the value of ψ can then be calculated. This form of the equations was presented by David & Johnson (1952); they suggested an iterative method, equivalent in the present notation to

$$k_{i+1} = [k_i \psi(\varpi_i, k_i)],$$

and provided a table of values of $p = (1-\varpi)/\varpi$ for values of ϕ (in their notation, y) 0·40 (0·05) 1·00. This process, however, appears to converge rather slowly, and a solution can, in many cases, be reached more expeditiously by a trial-and-error process, starting with the moment, or some other inefficient, estimate of k. For this value of k the value of ψ can be evaluated, together with the slope of the curve ($\psi = \psi(k)$) defined by (15) and (16):

$$\psi'(k) = \frac{1}{\phi-\varpi}[\psi - m\varpi^{k+1}]\left[\frac{\phi}{k} - \frac{k}{nm}\sum_{j=1}^{R}(k+j-1)^{-2}\sum_{r=j}^{R} n_r\right] - \frac{\psi}{k}[1-\phi(\varpi^k)]. \tag{17}$$

These values provide a second approximation to the root, which can then be found by a process of successive linear interpolation and extrapolation. Final adjustments may be made, if required, by the usual maximum-likelihood iterative procedure: this will usually be unnecessary, but the quantities (14) can be calculated, to provide an estimate of the asymptotic variance-covariance matrix of the estimates.

As with the moment method, extreme samples can occur for which the equations (15) and (16) have no solutions with $k > 0$. Unfortunately, the function $\psi(k)$ is of considerably more complicated a form than the function $\zeta(\varpi)$ occurring in the moment method, and the existence problem is correspondingly more difficult of solution. It can be shown that

$$\psi(0) = 1,$$

$$\psi'(0) = \frac{1}{n}\sum_{j=2}^{R}\frac{1}{(j-1)}\sum_{r=j}^{R} n_r + \tfrac{1}{2}\log \alpha,$$

where $$\phi(\alpha) = 1/m,$$

and that $$\lim_{k\to\infty} \psi(k) = m(1-e^{-\theta})/\theta,$$

where $$\theta = \frac{2}{nm}\sum_{j=2}^{R}(j-1)\sum_{r=j}^{R} n_r.$$

Clearly the conditions

$$\psi'(0) > 0, \quad \lim_{k\to\infty}\psi(k) < 1$$

are sufficient to ensure the existence of at least one solution, and I would conjecture that they are necessary, and also sufficient to ensure uniqueness, but have been unable to prove these results.

The question remains: What action is to be taken when the maximum-likelihood (or moment) method fails to give an acceptable solution? It seems reasonable to hope that one or other of the two limiting forms of the negative binomial, the Poisson distribution and Fisher's 'logarithmic series' distribution, will provide an adequate fit for all but some pathologically extreme (and highly improbable) samples which are unlikely to be fitted satisfactorily by any meaningful distribution.

9. An example of the maximum-likelihood calculations

The data are those already used in § 4; details of the calculations are shown in Table 3. The quantity Σn_r, used in the calculation of ϕ, is tabulated in the third column of Table 3 (*b*), and the remainder of the calculations are shown in some detail in Table 3 (*a*). In fact, this table shows far more detail than need be recorded in practice; if the calculations are made on a desk machine only the first three columns of Table 3 (*b*) and the successive values of k, ϖ and ψ need be written down. (The weighted sums of quotients in the second column of Table 3 (*a*) can be accumulated on the machine, and they, together with the values of ϕ, $\log\varpi$ and ϖ^k, are used as soon as they are calculated, and need not be written down.) However, ϕ and ϖ^k should be recorded for the first trial value, for use in evaluating ψ', and if final adjustments are to be made by the iterative process, all the quantities tabulated in Table 3 (*a*) should be recorded for the last trial value of k. The fourth and fifth columns of Table 3 (*b*) are recorded for the purpose of calculating $\Sigma(k+j-1)^{-2}\sum_{r=j}^{R} n_r$; the fourth, $(0{\cdot}63+j-1)^2$, for use in calculating ψ'; the fifth, $(0{\cdot}493+j-1)^2$, for use in calculating the variance-covariance matrix. The only remaining values required in this section are

$$nm = 110, \quad m = 3{\cdot}4375.$$

As an initial trial value the moment estimate of k, 0·633, was rounded to two decimal places, giving

$$\psi(0{\cdot}63) = 0{\cdot}9922.$$

As a guide to the location of the second trial value, ψ' was evaluated for $k = 0{\cdot}63$. Accumulating on the machine,

$$\sum_{j=1}(0{\cdot}63+j-1)^{-2}\sum_{r=j}^{R} n_r = 92{\cdot}2942,$$

whence

$$\psi' = \frac{(0{\cdot}9922-0{\cdot}3179)(0{\cdot}7008-0{\cdot}5286)}{0{\cdot}2094} - \frac{0{\cdot}9922\times 0{\cdot}3905}{0{\cdot}63}$$

$$= -0{\cdot}0605.$$

The second trial value was therefore taken as

$$0{\cdot}63-(1-0{\cdot}9922)/0{\cdot}0605 = 0{\cdot}50.$$

The fourth value, $k = 0{\cdot}493$, gives a value of ψ equal to 1 to four decimal places. There is no hope of improving further on this estimate by the trial-and-error method, without taking more decimal places, and using a more complicated interpolation formula, in the

Table 3. *Maximum-likelihood calculations for the example of* § 8

(*a*) Trial-and-error calculations

k	$\Sigma(k+j-1)^{-1}\sum_{r=j}^{R} n_r$	ϕ	ϖ	$\log_e \varpi$	ϖ^k	ψ	ψ'
0·63	77·0905	0·4415	0·2321	−1·46059	0·3984	0·9922	−0·0605
0·50	91·9239	0·4178	0·2124	−1·54930	0·4609	0·99952	
0·491	93·2185	0·4161	0·2110	−1·55590	0·4658	1·00016	
0·493	92·9270	0·4165	0·2113	−1·55450	0·4647	0·99996	

(*b*) Subsidiary tabulations

j	n_j	$\sum_{r=j}^{R} n_r$	$(j-0{\cdot}37)^2$	$(j-0{\cdot}507)^2$	Expected values of n_j			
					Truncated Poisson	Truncated negative binomial		Log series
						Moments	M.L.	
1	11	32	0·3969	0·243049	4·01	10·31	10·80	13·32
2	6	21	2·6569	2·229049	6·64	6·44	6·36	5·85
3	4	15	6·9169	6·215049	7·33	4·33	4·17	3·43
4	5	11	13·1769	12·201049	6·07	3·01	2·87	2·26
5	0	6	21·4369	20·187049	4·02	2·13	2·03	1·59
6	1	6	31·6969	30·173049	2·22	1·53	1·47	1·16
7	0	5	43·9569	42·159049	1·05	1·11	1·07	0·88
8	2	5	58·2169	56·145049	0·43	0·81	0·79	0·67
9	1	3	74·4769	72·131049	0·16	0·60	0·59	0·53
10	0	2	92·7369	90·117049	0·05	0·44	0·44	0·42
11	1	2	112·9969	110·103049				
12	0	1	135·2569	132·089049	0·02	1·29	1·41	1·89
13	1	1	159·5169	156·075049				

table of ϕ. There is, in fact, no great need of any further improvement, but, if desired, final increments to k and ϖ can be calculated using the variance-covariance matrix, which requires calculation in any case.

The fifth column of Table 3 (*b*) shows values of

$$(0{\cdot}493+j-1)^2,$$

and from these values (accumulating quotients on the machine)

$$\Sigma\,(0{\cdot}493+j-1)^{-2}\sum_{r=j}^{R} n_r = 145{\cdot}196464173.$$

Then, from this value and the entries in the last line of Table 3 (*a*), the estimate of the variance-covariance matrix, calculated from formulae (14), is

$$[V] = \begin{bmatrix} 554{\cdot}4186 & -94{\cdot}69298 \\ -94{\cdot}69298 & 19{\cdot}792976 \end{bmatrix}^{-1} = \begin{bmatrix} 0{\cdot}00986279 & 0{\cdot}0471853 \\ 0{\cdot}0471853 & 0{\cdot}276265 \end{bmatrix}.$$

Inserting the values $k = 0{\cdot}493$, $\varpi = 0{\cdot}2113$ into the left-hand sides of the maximum-likelihood equations (11) and (12), we obtain 0·006 and −0·0003 respectively. Thus the final increments are

$$\begin{bmatrix} \delta\varpi \\ \delta k \end{bmatrix} = [V] \begin{bmatrix} 0{\cdot}006 \\ -0{\cdot}0003 \end{bmatrix} = \begin{bmatrix} 0{\cdot}000045 \\ 0{\cdot}00020 \end{bmatrix}.$$

Table 4. *The function* $\phi(x) = \dfrac{-x \log_e x}{1-x}$

x	0·00	0·01	0·02	0·03	0·04	0·05	0·06	0·07	0·08	0·09
0·0	0	0·0465	0·0798	0·1085	0·1341	0·1577	0·1796	0·2002	0·2196	0·2381
0·1	0·2558	0·2728	0·2891	0·3049	0·3201	0·3348	0·3491	0·3629	0·3764	0·3896
0·2	0·4024	0·4149	0·4271	0·4390	0·4507	0·4621	0·4733	0·4843	0·4950	0·5056
0·3	0·5160	0·5262	0·5362	0·5461	0·5558	0·5653	0·5747	0·5839	0·5930	0·6020
0·4	0·6109	0·6196	0·6282	0·6367	0·6451	0·6533	0·6615	0·6695	0·6775	0·6854
0·5	0·6931	0·7008	0·7084	0·7159	0·7233	0·7307	0·7380	0·7451	0·7522	0·7593
0·6	0·7662	0·7731	0·7800	0·7867	0·7934	0·8000	0·8066	0·8131	0·8195	0·8259
0·7	0·8322	0·8385	0·8447	0·8509	0·8570	0·8630	0·8690	0·8750	0·8809	0·8868
0·8	0·8926	0·8983	0·9041	0·9097	0·9154	0·9209	0·9265	0·9320	0·9374	0·9429
0·9	0·9482	0·9536	0·9589	0·9642	0·9694	0·9746	0·9797	0·9848	0·9899	0·9950

The maximum-likelihood estimates are therefore

$$\hat{\varpi} = 0{\cdot}2113 \pm 0{\cdot}0993, \quad \hat{k} = 0{\cdot}493 \pm 0{\cdot}526.$$

The variance-covariance matrix for the moment estimates, recalculated in terms of the maximum-likelihood estimates to provide a more valid comparison, is

$$\begin{bmatrix} 0{\cdot}01364125 & 0{\cdot}0693156 \\ 0{\cdot}0693156 & 0{\cdot}405868 \end{bmatrix}.$$

The ratio of the determinants of the two matrices, converted to a percentage efficiency, is

$$100 \times 0{\cdot}000498/0{\cdot}000732 = 68{\cdot}1\,\%.$$

The efficiency is much the same for the two parameters, the actual values being 72·3 % for ϖ, and 68·1 % for k.

The maximum-likelihood estimate of k does not differ significantly from 0, which suggests that the data might be adequately fitted by Fisher's logarithmic series distribution. This, in fact, proves to be the case. The last four columns of Table 3 (*b*) show the expected numbers of cells with j breaks on the basis of four fitted distributions, a truncated Poisson (fitted by maximum likelihood), the two negative binomials fitted above, and the logarithmic series distribution (fitted by maximum likelihood). The fit of the Poisson distribution is obviously very poor; equally obviously the other three are all very good. The negative binomials

appear, superficially, to give a slightly better representation of the data, but this impression is largely due to the very good agreement in the first group. In fact the justification for using the negative binomial rather than the logarithmic series distribution comes, not from the data presented here, but from the whole series of experiments, only one of which is used here. In this series some distributions could be fitted adequately by the truncated Poisson, some by the logarithmic series distribution, and some not very satisfactorily by either, but the negative binomial form, with appropriate k, gave a good fit in all cases.

I am indebted to Dr C. E. Ford, of the Atomic Energy Research Establishment, Harwell, for permission to use his data in my examples, and to Miss A. D. Outhwaite, for drawing my attention to an omission from § 3.

REFERENCES

ANSCOMBE, F. J. (1950). Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika*, **37**, 358.

BLISS, C. I. & FISHER, R. A. (1953). Fitting the negative binomial distribution to biological data, with a note on the efficient fitting of the negative binomial. *Biometrics*, **9**, 176.

DAVID, F. N. & JOHNSON, N. L. (1952). The truncated Poisson. *Biometrics*, **8**, 275.

FISHER, R. A. (1941). The negative binomial distribution. *Ann. Eugen., Lond.*, **11**, 182.

GREENWOOD, M. & YULE, G. U. (1920). An enquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J. R. Statist. Soc.* **83**, 255.

HALDANE, J. B. S. (1941). The fitting of binomial distributions. *Ann. Eugen., Lond.*, **11**, 179.

THE RANDOMIZATION ANALYSIS OF A GENERALIZED RANDOMIZED BLOCK DESIGN†

By M. B. WILK
Iowa State College

1. Introduction

(1·1) *The experimental situation and design*

Suppose that t treatments are given whose properties (yields, responses, effects, etc.) we wish to compare when they interact with a given set of rs experimental units, the latter being classified into r blocks, each containing $s = pt$ units. Suppose, further, that an experiment is carried out in which the treatments are applied at random to the experimental units, with the restriction that each treatment appears with p units in each of the r blocks. We refer to this design as the generalized randomized block design and note that it includes as special cases the completely randomized design ($r = 1, p > 1$) and the randomized block design ($r > 1$, $p = 1$).

The object of this paper is to study the basis for statistical inference which is provided by the randomization procedure.

(1·2) *Some previous work*

The introduction of the device of randomization in the statistical design of experiments is due to R. A. Fisher (1926, 1935). A brief review of some important contributions to the problem of inference from randomized experiments is given by Wilk (1953).

In particular, we note here that the pattern of the present study leans heavily on the finite model analyses given by Kempthorne (1952*a*, *b*). The means and variances of the analysis of variance sums of squares obtained by Welch (1937) and Pitman (1937) for the randomized block design, and by Kempthorne (1952*b*) for the completely randomized design, and the expectations under randomization of the analysis of variance mean squares for randomized blocks with non-additivity given by Kempthorne (1952*a*), derive as special cases from the results (Wilk, 1953) on which the present paper is based.

(1·3) *Experimental error and randomization*

In many experimental situations it seems reasonable to distinguish two sources of experimental error, namely, the failure of different experimental units treated alike to respond identically, and the inability to reproduce an applied treatment exactly. The first of these, which we shall refer to as the unit error, stems from variation among the experimental units. The second type, which we shall call the technical error, stems from limitations on experimental technique.

Generally, the unit error is to be regarded as a fixed quantity associated with any given experimental unit; while the technical error may often be idealized as a random variable, say following a normal distribution with mean 0. The process of randomization is, usually, irrelevant to our conception of the technical errors, but is of critical importance in our treatment of the unit errors. It might in fact be said that the main function of randomization in experimental design is to control, in a statistical sense, the unit errors.

† Journal Paper No. J-2696 of the Iowa Agricultural Experiment Station, Ames, Iowa. Project No. 890.

To focus attention on the basis for statistical inference which is provided by the randomization procedure we shall in this paper assume that the only important source of experimental error is the unit errors.

2. THE ANALYSIS OF VARIANCE

(2·1) *The conceptual underlying population*

With the possible application of each of the treatments to each of the experimental units we associate a real (unknown) number. This defines a set of rst numbers which we take to be the conceptual underlying population for this experiment. Since the fact of the experimental situation is that each unit can be 'used' only once, the population defined is conceptual in the sense that only a subset of rs of the numbers of interest can be observed. The scope of a statistical inference for this situation can be delineated by noting that the conceivable totality of experimental information would be given by applying each treatment to every experimental unit and observing the response.

(2·2) *The population model*

Let $i = 1, 2, \ldots, r$ denote the block number.

Let $j = 1, 2, \ldots, s$ denote the experimental unit number within each block, where $s = pt$.

Let $k = 1, 2, \ldots, t$ denote the treatment number.

Let y_{ijk} represent the (conceptual) response which would be obtained if treatment k were applied to the jth unit in the ith block. Thus our underlying population is the set of (conceptual) unknown numbers $\{y_{ijk}\}$.

We will employ the usual dot convention for means, for example, $y_{..k} = \sum_{ij} y_{ijk}/(rs)$.

We now define

$$\begin{aligned} \mu &= y_{...}, \\ b_i &= y_{i..} - y_{...}, \\ t_k &= y_{..k} - y_{...}, \\ (bt)_{ik} &= (y_{i.k} - y_{i..}) - (y_{..k} - y_{...}), \\ e_{ij} &= y_{ij.} - y_{i..}, \\ n_{ijk} &= (y_{ijk} - y_{ij.}) - (y_{i.k} - y_{i..}), \end{aligned}$$

where $i = 1, 2, \ldots, r$; $j = 1, 2, \ldots, s$; $k = 1, 2, \ldots, t$.

These quantities may be given a physical interpretation.

μ is the (conceptual) overall mean yield which would be obtained if each treatment were applied to every unit in every block.

b_i is the difference between the (conceptual) mean yield of all treatments on all units of block i and μ, and may be thought of as the effect attributable to the ith block.

In an analogous way t_k may be thought of as the effect attributable to the kth treatment. We note that by definition this is the average effect over all experimental units.

$(bt)_{ik}$ is the difference between the effect of treatment k on block i and t_k. Thus it is a measure of the extent to which treatment k and block i interact and will be called the block-treatment interaction.

e_{ij} gives the difference between the (conceptual) mean of the yields of all treatments on unit j of block i and the mean over the whole block. It, therefore, measures the extent to which the jth unit deviates from the other units of block i and will be called the unit error.

The set $\{e_{ij}\}, j = 1, 2, \ldots, s$, can be used to give a measure of the heterogeneity of units within the ith block.

n_{ijk} is the difference between the effect of treatment k on unit j of block i and the effect of treatment k over all of block i. It is, therefore, a measure of the extent to which treatment k and unit j of block i interact and will be called the unit-treatment interaction.

The following equation is an algebraic indentity:

$$y_{ijk} = \mu + b_i + t_k + (bt)_{ik} + e_{ij} + n_{ijk}.$$

From the definition of the quantities it follows that

$$\sum_i b_i = \sum_k t_k = \sum_i (bt)_{ik} = \sum_k (bt)_{ik} = \sum_j e_{ij} = \sum_j n_{ijk} = \sum_k n_{ijk} = 0.$$

Algebraically the sets of numbers $\{b_i\}$, $\{t_k\}$, $\{(bt)_{ik}\}$, $\{e_{ij}\}$ and $\{n_{ijk}\}$ are pairwise independent in the sense that, for example, $t_k = 0$ for all k does not imply $(bt)_{ik} = 0$ for all k. In particular, if $(bt)_{ik} = 0$ for all i and k, then algebraically this does not imply that $n_{ijk} = 0$ for all i and k. On the other hand, the physical situation is such that if all the block-treatment interactions were zero, then one would expect that all the unit-treatment interactions would be zero. This follows from the fact that the experimental units are blocked so as to be more homogeneous within a block than from block to block. If so, then the lack of a differential effect from block to block would lead one to expect that the differential effect from unit to unit within a block would be negligible.

The converse is, of course, not true. If the units within a block are sufficiently homogeneous then the unit-treatment interactions will be small in absolute value. But this does not preclude important differences between blocks and hence the possible existence of non-additive effects from block to block, i.e. block-treatment interactions. It would appear therefore that if the blocking of experimental units is successful, then the n_{ijk} will be negligible while the $(bt)_{ik}$ may be important.

The essential point here is that it may not be unrealistic to assume that the treatments react additively (i.e. unit-treatment interactions are zero) within a block even though they react non-additively from block to block.

In many instances, whether we assume the unit-treatment interactions are zero or not, it may be reasonable to assume that the variability of units within a block is essentially the same for all blocks. This assumption could be idealized as

$$\sum_j e_{ij}^2 = (s-1)\,\sigma^2;$$

$$\sum_j n_{ijk}^2 = (s-1)\,\sigma_n^2 \quad (i = 1, 2, \ldots, r;\ k = 1, 2, \ldots, t).$$

The importance of this discussion is in indicating the direction of simplifying assumptions in the analysis of the design.

(2·3) *The statistical model*

In actually carrying out the experiment, we will in fact observe only a (restricted) random sample of size rs from the set of rst numbers $\{y_{ijk}\}$. Let x_{ikf} denote the observation obtained from the fth replication of the kth treatment in the ith block, where $i = 1, 2, \ldots, r$; $k = 1, 2, \ldots, t$; and $f = 1, 2, \ldots, p$ for each (ik). Thus $\sum_f x_{ikf}$ represents the observed total response from all units in block i to which treatment k has been applied.

To write an explicit model for $\sum_j x_{ikf}$ we now define some additional quantities. Let†

$$D_{ij}^k = 1 \quad \text{if treatment } k \text{ falls on unit } j \text{ of block } i$$
$$= 0 \quad \text{otherwise.}$$

Because random methods of allocation are employed, the D_{ij}^k may be treated as random variables, and from the design of the experiment it is easy to specify certain of the distributional properties of the D_{ij}^k. For example,

$$P(D_{ij}^k = 1) = P \text{ (treatment } k \text{ falls on unit } j \text{ of block } i)$$
$$= p/s.$$

Hence
$$E(D_{ij}^k) = p/s,$$

where we follow usual convention and define $E(\alpha)$ to be the mathematical expectation of α. More detail on the D_{ij}^k is given by Wilk (1953).

It is easy to see that

$$\sum_f x_{ikf} = p[\mu + b_i + t_k + (bt)_{ik}] + \sum_j e_{ij} D_{ij}^k + \sum_j n_{ijk} D_{ij}^k.$$

This relation we shall call the statistical model. This formulation exhibits explicitly just what are the random variables involved in the model, namely, the D_{ij}^k which take on the values 0 or 1 with known probabilities. We note that it is the physical act of randomization in allocation of treatments to units that permits us to treat the D_{ij}^k as random variables, and hence provides some basis for statistical inference.

(2·4) *The analysis of variance table*

The primitive analysis of variance for this design is simply a breakdown of the sum of squares of deviations of individual observations from their mean into additive components which can be attributed to various sources. The analysis of variance has proved useful in the statistical analysis of experiments in the estimation of components of variation, in the estimation of the variance of estimates of treatment comparisons, and in making tests of significance.

As before, we use the dot convention for means, e.g. $x_{i\cdot\cdot} = \sum_{kf} x_{ikf}/s$. The algebraic detail of the analysis of variance is given in Table 1.

We can now employ the statistical model for $\sum_f x_{ikf}$, and certain properties of the D_{ij}^k, to derive the expectations under randomization of the analysis of variance mean squares. The detailed algebra is given by Wilk (1953).

The results are tabulated in Table 2.

It is of interest to note that if the n_{ijk} are not all zero, then even if $t_k = 0$ for all k, the expectation of the treatment mean square is not equal to the expectation of the error mean square. Similarly for the interaction mean square. If the n_{ijk} are small compared with the e_{ij}, or if t is large, the bias is negligible. The assumptions

$$\sum_j e_{ij}^2 = (s-1)\sigma^2, \quad \text{all } i,$$
$$\sum_j n_{ijk}^2 = (s-1)\sigma_n^2, \quad \text{all } i \text{ and } k,$$

do not affect the above discussion.

† A more formal definition of the D_{ij}^k, in which they appear as characteristic set functions of set-valued random variables, is given by Wilk (1953).

Clearly, for small t, a meaningful comparison of the treatment mean square (or interaction mean square) with the error mean square depends heavily on the assumption that the n_{ijk} are negligible.

Table 1. *Analysis of variance*

Due to	Degrees of freedom	Sum of squares	Mean squares
Blocks	$(r-1)$	$B = s\sum_i (x_{i..}-x_{...})^2$	$B^* = \dfrac{B}{r-1}$
Treatments	$(t-1)$	$T = rp\sum_k (x_{.k.}-x_{...})^2$	$T^* = \dfrac{T}{(t-1)}$
Interactions	$(r-1)(t-1)$	$I = p\sum_{ik} (x_{ik.}-x_{i..}-x_{.k.}+x_{...})^2$	$I^* = \dfrac{I}{(r-1)(t-1)}$
Error	$rt(p-1)$	$R = \sum_{ikf} (x_{ikf}-x_{ik.})^2$	$R^* = \dfrac{R}{rt(p-1)} = v$
Total	$rs-1$	$G = \sum_{ikf} (x_{ikf}-x_{...})^2$	

Table 2. *Expectations of mean squares under randomization*

Due to	Mean square	Expectation of mean square
Blocks	B^*	$\dfrac{1}{rt(s-1)}\sum_{ijk} n_{ijk}^2 + \dfrac{pt}{(r-1)}\sum_i b_i^2$
Treatments	T^*	$\dfrac{1}{r(s-1)}\sum_{ij} e_{ij}^2 + \dfrac{(t-2)}{rt(s-1)(t-1)}\sum_{ijk} n_{ijk}^2 + \dfrac{rp}{(t-1)}\sum_{ik} (bt)_{ik}^2$
Interactions	I^*	$\dfrac{1}{r(s-1)}\sum_{ij} e_{ij}^2 + \dfrac{(t-2)}{rt(s-1)(t-1)}\sum_{ijk} n_{ijk}^2 + \dfrac{p}{(r-1)(t-1)}\sum_{ik} (bt)_{ik}^2$
Error	R^*	$\dfrac{1}{r(s-1)}\sum_{ij} e_{ij}^2 + \dfrac{1}{rt(s-1)}\sum_{ijk} n_{ijk}^2$

In contrast with similar results based on normal theory, the expectation of the mean square for blocks does not contain all components which appear in the expectation of the error mean square. The same remark applies to a comparison of blocks mean square with interaction mean square when all $(bt)_{ik} = 0$. It is apparent that an analysis of variance test of significance of block effects, for the situation under consideration, cannot be justified by randomization.

(2·5) *Some randomization moments*

In Table 3 we give some moments under randomization of the analysis of variance sums of squares, under the hypothesis that $t_k = (bt)_{ik} = 0$ for all i and k, and using the simplifying assumptions:

$$n_{ijk} = 0, \qquad \text{for all } i, j, k,$$

$$\sum_j e_{ij}^2 = (s-1)\sigma^2, \qquad \text{for all } i,$$

$$\sum_j e_{ij}^4 = (s-1)(2s-3)\sigma^4/s, \qquad \text{for all } i.$$

We use G_0, B_0, T_0, I_0 and R_0 to denote the values of G, B, T, I and R, respectively, under the indicated conditions.

The results of Table 3 are derived from more general expressions, given by Wilk (1953), which use no homogeneity assumptions regarding $\sum_j e_{ij}^2$ and $\sum_j e_{ij}^4$. The algebraic development is also detailed in the same paper.

Table 3. *Some randomization moments under simplifying assumptions*

$$E(G_0) = G_0 = pt\sum_i b_i^2 + r(s-1)\sigma^2 \qquad V(G_0) = 0$$

$$E(B_0) = B_0 = pt\sum_i b_i^2 \qquad V(B_0) = 0$$

$$E(T_0) = (t-1)\sigma^2 \qquad V(T_0) = 2(t-1)\sigma^4\left(1-\frac{1}{rp}\right)$$

$$E(I_0) = (r-1)(t-1)\sigma^2 \qquad V(I_0) = 2(r-1)(t-1)\sigma^4\left(1-\frac{r-1}{rp}\right)$$

$$E(R_0) = rt(p-1)\sigma^2 \qquad V(R_0) = \frac{2r(t-1)(p-1)}{p}\sigma^4$$

$$\operatorname{cov}(T_0, I_0) = -\frac{2(t-1)(r-1)}{rp}\sigma^4 \qquad \operatorname{cov}(T_0, R_0) = -\frac{2(p-1)(t-1)}{p}\sigma^4$$

$$\operatorname{cov}(I_0, R_0) = -\frac{2(p-1)(t-1)(r-1)}{p}\sigma^4$$

It is of some interest to note that for large r or p, $V(T_0)$ approaches $2(t-1)\sigma^4$; for large p, $V(I_0)$ approaches $2(r-1)(t-1)\sigma^4$; for large p, $V(R_0)$ becomes independent of p and approaches $2r(t-1)\sigma^4$; if $p = 1$, $V(I_0)$ becomes $(2r-1)(t-1)\sigma^4/r$, and thence for large r becomes independent of r. The correspondence to and divergence from the corresponding moments based on normal theory will be apparent.

The correlation between T_0 and I_0 is $-\sqrt{\dfrac{(r-1)}{(rp-1)(rp-r+1)}}$, which is -1 for $p = 1$, and goes to 0 as p increases. The correlation between T_0 and R_0 is $-\sqrt{\dfrac{p-1}{rp-1}}$, which for large p goes to $-\sqrt{\dfrac{1}{r}}$, and for large r goes to zero. The correlation between I_0 and R_0 is

$$-\sqrt{\frac{(r-1)(p-1)}{(rp-r-1)}},$$

which for large p goes to $-\sqrt{\dfrac{(r-1)}{r}}$, and for large r goes to -1.

3. ESTIMATION

This section deals with the estimation of certain functions of the parameters of the population model. Variances of the estimates under various conditions are given, and the estimation of these variances is considered. Some additional results, as well as the algebraic detail, are given by Wilk (1953).

The following general notation is used:

$V(\alpha)$ is the variance of α under no assumptions.

$V_1(\alpha)$ is the variance of α under the homogeneity assumptions that

$$\sum_j e_{ij}^2 = (s-1)\sigma^2, \quad \sum_j n_{ijk}^2 = (s-1)\sigma_n^2 \quad \text{and} \quad \sum_j e_{ij} n_{ijk} = 0,$$

for each $i = 1, 2, \ldots, r$ and $k = 1, 2, \ldots, t$.

$V^0(\alpha)$ is the variance of α under the assumption that all the n_{ijk} are zero.

v denotes the analysis of variance error mean square.

(*a*) An unbiased estimate of μ is $\hat{\mu} = x_{...}$:

$V(\hat{\mu}) = \dfrac{1}{r^2ts(s-1)} \sum_{ijk} n_{ijk}^2$, and there appears to be no reasonable way to estimate $\sum_{ijk} n_{ijk}^2$.

$V^0(\hat{\mu}) = 0$, so μ is known without error if the n_{ijk} are all zero.

(*b*) An unbiased estimate of b_i is $\hat{b}_i = x_{i..} - x_{...}$:

$$V(\hat{b}_i) = \frac{(r-2)}{rts(s-1)} \sum_{jk} n_{ijk}^2 + \frac{1}{r^2ts(s-1)} \sum_{ijk} n_{ijk}^2.$$

$V_1(\hat{b}_i) = \dfrac{r-1}{rs}\sigma_n^2$. No reasonable estimate of either $V(\hat{b}_i)$ or $V_1(\hat{b}_i)$ is available.

$V^0(\hat{b}_j) = 0$, so the block effects are known precisely if all the n_{ijk} are zero.

(*c*) An unbiased estimate of t_k is $\hat{t}_k = x_{..k} - x_{...}$:

$$V(\hat{t}_k) = \frac{1}{r^2s(s-1)} \left[(t-1) \sum_{ij} e_{ij}^2 + 2(t-2) \sum_{ij} e_{ij} n_{ijk} + (t-3) \sum_{ij} n_{ijk}^2 + \frac{1}{t} \sum_{ijk} n_{ijk}^2 \right].$$

$V_1(\hat{t}_k) = \dfrac{(t-1)}{rs}(\sigma^2 + \sigma_n^2) - \dfrac{1}{rs}\sigma_n^2$, and $\dfrac{(t-1)}{rs} v$ tends to overestimate $V_1(\hat{t}_k)$.

$V^0(\hat{t}_k) = \dfrac{(t-1)}{r^2s(s-1)} \sum_{ij} e_{ij}^2$, and $\dfrac{(t-1)}{rs} v$ is an unbiased estimate of $V^0(\hat{t}_k)$.

(*d*) An unbiased estimate of $(bt)_{ik}$ is $\widehat{(bt)}_{ik} = x_{ik.} - x_{i..} - x_{..k} + x_{...}$:

$$\begin{aligned} V(\widehat{bt})_{ik} = \frac{1}{rs(s-1)} \Bigg[& (t-1)(r-2) \sum_j e_{ij}^2 + \frac{(t-1)}{r} \sum_{ij} e_{ij}^2 \\ & + 2(t-2)(r-2) \sum_j e_{ij} n_{ijk} + \frac{2(rt-r-2t+1)}{r} \sum_{ij} e_{ij} n_{ijk} \\ & + (rt-3r-2t+4) \sum_j n_{ijk}^2 + \frac{(t-3)}{r} \sum_{ij} n_{ijk}^2 \Bigg]. \end{aligned}$$

$V_1(\widehat{bt})_{ik} = \dfrac{(r-1)(t-1)}{rs}(\sigma^2 + \sigma_n^2) - \dfrac{r-1}{rs}\sigma_n^2$, and so $\dfrac{(r-1)(t-1)}{rs} v$ tends to overestimate $V_1(\widehat{bt})_{ik}$.

(*e*) An unbiased estimate of the treatment mean $(\mu + t_k)$ is $(\hat{\mu} + \hat{t}_k) = x_{..k}$:

$$V(\hat{\mu} + \hat{t}_k) = \frac{(t-1)}{r^2s(s-1)} - \sum_{ij} (e_{ij}^2 + 2e_{ij} n_{ijk} + n_{ijk}^2).$$

$V_1(\hat{\mu} + \hat{t}_k) = \dfrac{(t-1)}{rs}(\sigma^2 + \sigma_n^2)$, and $\dfrac{(t-1)}{rs} v$ is an unbiased estimate of $V_1(\hat{\mu} + \hat{t}_k)$.

$V^0(\hat{\mu} + \hat{t}_k) = \dfrac{(t-1)}{r^2s(s-1)} \sum_{ij} e_{ij}^2$, and $\dfrac{(t-1)}{rs} v$ is an unbiased estimate of $V^0(\hat{\mu} + \hat{t}_k)$.

(*f*) An unbiased estimate of the treatment contrast $\sum_k \rho_k t_k$, with $\sum_k \rho_k = 0$, is

$$\sum_k \rho_k \hat{t}_k = \sum_k \rho_k (x_{.k.} - x_{...}):$$

$$V(\sum_k \rho_k \hat{t}_k) = \frac{1}{r^2 p(s-1)} \left[(\sum_k \rho_k^2) \sum_{ij} e_{ij}^2 + 2 \sum_{ijk} \rho_k^2 e_{ij} n_{ijk} + \sum_{ijk} \rho_k^2 n_{ijk}^2 - \frac{1}{t} \sum_k (\sum_{ij} \rho_k n_{ijk})^2 \right].$$

$V_1(\sum_k \rho_k \hat{t}_k) = \frac{1}{rp} (\sum_k \rho_k^2)(\sigma^2 + \sigma_n^2) - \frac{1}{r^2 s(s-1)} \sum_{ij} (\sum_k \rho_k n_{ijk})^2$, and so $\frac{1}{rp} (\sum_k \rho_k^2) v$ tends to overestimate $V_1(\sum_k \rho_k \hat{t}_k)$.

$V^0(\sum_k \rho_k \hat{t}_k) = \frac{1}{r^2 p(s-1)} (\sum_k \rho_k^2) \sum_{ij} e_{ij}^2$, and this is estimated unbiasedly by $\frac{1}{rp} (\sum_k \rho_k^2) v$.

We close this section with a brief discussion of special cases. If $r = 1$, $p > 1$ (completely randomized design), the discussion of block effects and interactions is to be ignored, and the remaining results carry over directly.

For the case of $p = 1$, $r > 1$ (randomized block design) the situation is somewhat different in that the analysis of variance error mean square which we have been using becomes non-existent. If we can assume that all block-treatment interactions are zero, then the interaction mean square in the randomized block design has expectation

$$\frac{1}{r(t-1)} \sum_{ij} e_{ij}^2 + \frac{(t-2)}{rt(t-1)^2} \sum_{ijk} n_{ijk}^2,$$

and for most of the cases discussed above, the variances of estimates may be estimated unbiasedly.

Kempthorne (1952*a*, *b*) has discussed in detail the randomization analysis of the randomized block and completely randomized designs.

4. TESTS OF SIGNIFICANCE

In this section we consider tests of significance of a number of null hypotheses (hypotheses of equivalence) of possible interest. The object of the test is to obtain a measure of the adequacy of the experiment to indicate conclusions.

To employ a randomization test of a null hypothesis we require no assumptions about the form of the frequency distribution of the observations. A real-valued function of the observations is selected which will reflect (in a monotone increasing fashion, say) the deviation from equivalence of the treatments.

Having obtained a set of numbers for the particular experimental arrangement, we associate these numbers with the experimental units from which they derived. We then evaluate the function selected for the actual disposition of the treatments and for all other dispositions possible under the restriction of the design. A measure of the strength of the evidence against the null hypothesis (i.e. the level of significance† of the experiment) is the proportion of the values of the function which exceed the 'observed' value. The initial introduction of randomization enables a probabilistic interpretation.

In general, the amount of computation implied by the above procedure is prohibitive. Consequently, several writers (Welch, 1937; Pitman, 1937; Kempthorne, 1952*a*, *b*) have studied the possible approximation to the randomization test of certain procedures which

† A high level of significance corresponds to a low proportion.

derive from normal theory assumptions. Since the normal theory criteria fulfil the requirement for a randomization test function, the approach has been to examine the correspondence by comparing randomization means and variances of these criteria with their normal theory analogues. We proceed to make a similar examination for several null hypotheses of possible interest in the generalized randomized block design. The normal theory analysis for this design has been outlined by Wilk (1953).

In all that follows in this section we make the assumption that the n_{ijk} are all zero, and that $\sum_j e_{ij}^2 = (s-1)\sigma^2$, $\sum_j e_{ij}^4 = (s-1)(2s-3)\sigma^4/s$ $(i = 1, 2, \ldots, r)$.

(*a*) Consider the null hypothesis that the treatments all react identically on every experimental unit. This implies that $y_{ijk} = y_{ij.}$ for all i, j, k, and hence $t_k = (bt)_{ijk} = n_{ijk} = 0$ for all i, j, k.

The normal theory test of this hypothesis may be based on the criterion

$$(T_0 + I_0)/(T_0 + I_0 + R_0),$$

which under the usual normal theory assumptions and under the present null hypothesis follows a Beta distribution

$$f(q) = Kq^{\frac{1}{2}f_1 - 1}(1-q)^{\frac{1}{2}f_2 - 1} \quad (0 < q < 1),$$

where $f_1 = r(t-1)$, $f_2 = rt(p-1)$.

Let U and U^* denote the criterion above under normal theory and randomization respectively. Then $O < U < 1$ and $O < U^* < 1$. Since the denominator of U^* is constant under randomization $E(U^*) = \dfrac{(t-1)}{(s-1)}$ and $V(U^*) = \dfrac{2(t-1)(p-1)}{rp(s-1)^2}$. Under normal theory, $E(U) = \dfrac{(t-1)}{(s-1)}$, $V(U) = \dfrac{2t(t-1)(p-1)}{(rs-r+2)(s-1)^2}$. Thus $E(U^*) = E(U)$; and for large s or small r, $V(U)$ approaches $V(U^*)$.

If we can judge the correspondence of the distributions of U and U^* from the comparison of their ranges, means and variances, it would appear that the randomization test may be approximated by the usual test based on normal theory.

To better the approximation to the randomization distribution of U^* by a Beta distribution, we might use a Beta variate Q, with density

$$f(Q) = KQ^{\frac{1}{2}f_1 - 1}(1-Q)^{\frac{1}{2}f_2 - 1} \quad (0 < Q < 1),$$

such that $$E(Q) = \frac{f_1}{f_1 + f_2} = \frac{(t-1)}{(s-1)} = E(U^*),$$

and $$V(Q) = \frac{2f_1 f_2}{(f_1+f_2)^2(f_1+f_2+2)} = \frac{2(t-1)(p-1)}{rp(s-1)^2} = V(U^*).$$

Then $$f_1 = \frac{(t-1)(rs-2)}{(s-1)}, \quad f_2 = \frac{t(p-1)(rs-2)}{(s-1)}.$$

For different assumptions regarding the quantities $\sum_j e_{ij}^2$ and $\sum_j e_{ij}^4$ $(i = 1, 2, \ldots, r)$, we would thus arrive at different Beta distributions to approximate the randomization test. Welch (1937) considered such an adjustment for the case of the randomized block design when the blocks do not exhibit equal variation.

(*b*) We consider now a situation in which we can assume that $(bt)_{ik} = 0$ for all i and k, and wish to test the null hypothesis that $t_k = 0, k-1, 2, \ldots, t$.

The criterion suggested by normal theory for this situation is $T_0/(T_0+I_0+R_0)$. Let W denote this quantity under normal theory, and W^* denote it under randomization.

Then under the null hypothesis, $O < W < 1$, $O < W^* < 1$, and W is a Beta variate with $E(W) = \dfrac{(t-1)}{r(s-1)}$ and $V(W) = \dfrac{2(t-1)(rs-r-t-1)}{r^2(s-1)^2(rs-r-2)}$; while $E(W^*) = \dfrac{(t-1)}{r(s-1)}$ and

$$V(W^*) = \frac{2(t-1)}{r^2(s-1)^2}\left(1-\frac{1}{rp}\right).$$

Thus $E(W) = E(W^*)$, and for large values of p or r $V(W)$ and $V(W^*)$ are approximately equal.

Of course as in (*a*) we can define a Beta variate which will have the same mean and variance as W^* and, perhaps, in this way better the approximation.

A special case of interest is that in which $p = 1, r > 1$, i.e. the randomized block design. In that case, R does not exist, and carrying out a meaningful analysis of variance test of significance for that design hinges upon the assumption that the block-treatment interactions are zero.

Also, of particular concern here is the case of $r = 1, p > 1$, i.e. the completely randomized design. In that case, I does not exist, and the test we have discussed is the one of interest for that design.

It is a pleasure to acknowledge the assistance of Prof. O. Kempthorne in the preparation of this paper.

REFERENCES

FISHER, R. A. (1926). The arrangement of field experiments. *J. Minist. Agric.* **33**, 503–13.
FISHER, R. A. (1935). *The Design of Experiments*. Edinburgh: Oliver and Boyd.
KEMPTHORNE, O. (1952*a*). *The Design and Analysis of Experiments*. New York: John Wiley and Sons.
KEMPTHORNE, O. (1952*b*). The randomization theory of experimental inference. Paper delivered to the Amer. Stat. Ass. and Biometrics Soc. Publication pending.
PITMAN, E. J. G. (1937). Significance tests which may be applied to samples from any population. III. The analysis of variance test. *Biometrika*, **29**, 322–35.
WELCH, B. L. (1937). On the *z*-test in randomized blocks and latin squares. *Biometrika*, **29**, 21–52.
WILK, M. B. (1953). The randomization analysis of a generalized randomized block design. M.S. thesis, Department of Statistics, Iowa State College.

SOME QUICK SIGN TESTS FOR TREND IN LOCATION AND DISPERSION

By D. R. COX

Statistical Laboratory, University of Cambridge

and A. STUART

Division of Research Techniques, London School of Economics

1. Introduction and summary

Many distribution-free tests have been devised to test the hypothesis of randomness of a series of N observations, i.e. the hypothesis that N independent random variables have the same continuous distribution function. Of these, the rank correlation tests are the most efficient tests against normal trend alternatives, but others are of some use in situations where speed and simplicity of computation are important.

In this paper, we discuss a class of simple sign tests, considered first as tests against trend in location. Optimum tests are found from the standpoint of asymptotic relative efficiency (a measure of local power in large samples), and it appears that the best of these tests may be preferred to the other simple tests considered in the literature, although they are, of course, less efficient than the rank correlation tests.

Similar tests are available for trend in dispersion, and the efficiency of these, in the normal situation, is investigated and compared with the test based on the maximum-likelihood estimator. Finally, we add a few remarks on sequential sign tests.

Readers not interested in the theory should look at §§ 10, 11 and 14, where there are brief statements of the tests and numerical examples.

2. The sign tests for trend in location

We consider a series of N independent observations from a standardized normal regression model with an upward trend, i.e.

$$H_1: \quad y_i = \alpha + \Delta i + \epsilon_i \quad (i = 1, 2, \ldots, N),$$

where $\Delta \geqslant 0$ and the ϵ_i are independent standardized normal variates. We wish to test the null hypothesis

$$H_0: \quad \Delta = 0,$$

using a distribution-free test statistic, so that our test will remain valid whatever the continuous distribution of the ϵ terms in the model, although naturally its efficiency will vary with the form of the distribution.

The most efficient known distribution-free tests of H_0 are those based on the rank correlation coefficients (Stuart, 1954), but our object here is specifically to find tests which are quick and simple to compute. We define for $i < j$ the score

$$h_{ij} = \begin{cases} +1 & \text{if} \quad y_i > y_j, \\ 0 & \text{if} \quad y_i < y_j. \end{cases}$$

h_{ij} is thus based on a comparison of the ith and jth in the series of observations. The distribution of the observations will be assumed continuous so that the possibility of ties can be ignored (see, however, § 16).

We confine ourselves throughout to comparisons of *independent* pairs of observations, i.e. no observation is compared with more than one other observation. (This is in contrast to the procedure used in calculating the rank correlation coefficients, where every observation is compared with every other observation in the series.) Since there are N observations, there can be no more than $\frac{1}{2}N$ such independent comparisons. We now assume N to be even and always take $i<j$. Our problem is to find the set of comparisons and the appropriate weights w_{ij} which will make the statistic

$$S = \Sigma w_{ij} h_{ij} \tag{1}$$

as efficient a test of H_0 as possible. The summation in (1) contains $\frac{1}{2}N$ terms, no suffix being repeated. All tests of the form (1) are distribution-free, since on the null hypothesis any h_{ij} is a $0-1$ variate with probabilities $(\frac{1}{2}, \frac{1}{2})$, whatever the distribution of the ϵ_i.

3. ASYMPTOTIC RELATIVE EFFICIENCY

We shall use as our criterion of efficiency the asymptotic relative efficiency (A.R.E.) of a test. If there are two consistent tests, s and t, of a hypothesis $H_0: \Delta = 0$, the A.R.E. is the reciprocal of the ratio of sample sizes required to attain the same power against the same alternative hypothesis H_1, taking the limit as the sample size N tends to infinity and as H_1 tends to H_0. (This second limiting process is necessary to keep the power of consistent tests bounded away from 1.) Pitman (1948) and Noether (1955) have shown that, if s and t both have normal limiting distributions on H_0 and H_1, the A.R.E. of s compared to t is given by

$$\text{A.R.E.}\,(s,t) = \lim_{N\to\infty}\left(\frac{R^2(s)}{R^2(t)}\right)^{1/r}, \tag{2}$$

where

$$R^2(X) = \left\{\left[\frac{\partial}{\partial\Delta}E(X\,|\,\Delta)\right]_{\Delta=0}\right\}^2 \Big/ D^2(X\,|\,\Delta=0), \tag{3}$$

provided that r satisfies the equations

$$\lim_{N\to\infty} R^2(s)N^{-r} = R_1, \quad \lim_{N\to\infty} R^2(t)N^{-r} = R_2. \tag{4}$$

Here E and D^2 denote mean and variance as usual, while R_1 and R_2 are constants independent of N. The interpretation of the A.R.E. is discussed critically in § 9 below.

4. THE BEST SIGN TEST

Since h_{ij} is a 0–1 variate,

$$E(h_{ij}) = \text{prob}\,(y_i > y_j),$$

and as $(y_j - y_i)$ is a normal variate with mean $(j-i)\Delta$ and variance 2 this is

$$E(h_{ij}) = G\left\{-\frac{(j-i)\Delta}{\sqrt{2}}\right\}, \tag{5}$$

where

$$G(x) = \int_{-\infty}^{x}\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}t^2}dt.$$

Now

$$\left[\frac{\partial}{\partial x}G(x)\right]_{x=0} = \frac{1}{\sqrt{(2\pi)}}, \tag{6}$$

so that, by (5) and (6),

$$E'(h_{ij}) = \left[\frac{\partial}{\partial \Delta} E(h_{ij})\right]_{\Delta=0} = \frac{(i-j)}{2\sqrt{\pi}}. \tag{7}$$

We now write $(j-i) = r_{ij}$. Using (7) in (1), we obtain

$$E'(S) = \Sigma w_{ij} E'(h_{ij}) = -\frac{1}{2\sqrt{\pi}} \Sigma w_{ij} r_{ij}. \tag{8}$$

We also have

$$D^2(S \mid \Delta = 0) = \Sigma w_{ij}^2 V(h_{ij} \mid H_0) = \tfrac{1}{4} \Sigma w_{ij}^2. \tag{9}$$

Equations (3), (8) and (9) give

$$R^2(S) = \frac{1}{\pi} \frac{(\Sigma w_{ij} r_{ij})^2}{\Sigma w_{ij}^2}, \tag{10}$$

and we now wish to maximize (10) to obtain the highest possible A.R.E. We do this in two stages. First we maximize (10) with respect to the w_{ij}, regarding the r_{ij} as fixed, and then we choose the supremum of these maxima for variations in the r_{ij}.

To maximize (10) for fixed r_{ij} and variation in the w_{ij}, we must maximize $\Sigma w_{ij} r_{ij}$ subject to Σw_{ij}^2 being held constant, i.e. we must unconditionally maximize

$$F = \Sigma w_{ij} r_{ij} - \lambda \Sigma w_{ij}^2.$$

It is clear from the conditions of the problem that each w_{ij} will be a function of the corresponding r_{ij}, so that on differentiating F for a stationary value we get

$$r_{ij} + w_{ij} \frac{\partial r_{ij}}{\partial w_{ij}} - 2\lambda w_{ij} = 0,$$

i.e.

$$\frac{r_{ij}}{w_{ij}} + \frac{\partial r_{ij}}{\partial w_{ij}} = 2\lambda.$$

This is satisfied by

$$w_{ij} = \lambda r_{ij}, \tag{11}$$

so that the required set of weights are proportional to the distances apart of the observations compared. The stationary value is a maximum. Substituting (11) into (10), we have

$$R^2(S) = \frac{1}{\pi} \Sigma r_{ij}^2. \tag{12}$$

This is the maximum value of $R^2(S)$ for a fixed set of r_{ij}. The r_{ij} are a set of $\frac{1}{2}N$ differences between pairs of integers chosen from the integers $1, 2, \ldots, N$. It is easily seen that Σr_{ij}^2 is largest when the pairs are $(1, N)$, $(2, N-1)$, $(3, N-2)$ and so on. In general

$$r_{ij} = (N-k+1) - k = N - 2k + 1 \quad (k = 1, 2, \ldots, \tfrac{1}{2}N) \tag{13}$$

so that

$$\Sigma r_{ij}^2 = \sum_{k=1}^{\frac{1}{2}N} (N-2k+1)^2 = \tfrac{1}{6}N(N^2-1), \tag{14}$$

and the supremum value of (12) is therefore

$$R^2(S_1) = \frac{N(N^2-1)}{6\pi} \sim \frac{N^3}{6\pi}. \tag{15}$$

We have denoted by S_1 the optimum S statistic

$$S_1 = \sum_{k=1}^{\frac{1}{2}N} (N-2k+1) h_{k, N-k+1},$$

for which

$$\left.\begin{aligned} E(S_1 \mid \Delta = 0) &= \tfrac{1}{2}\Sigma(N-2k+1) = \tfrac{1}{8}N^2, \\ D^2(S_1 \mid \Delta = 0) &= \tfrac{1}{4}\Sigma(N-2k+1)^2 = \tfrac{1}{24}N(N^2-1). \end{aligned}\right\} \tag{16}$$

The test based on S_1 is essentially a simplified version of Spearman's rank correlation test, which is in effect defined by

$$V = \sum_{i<j}(j-i)\,h_{ij}, \tag{17}$$

where the summation extends over all possible $\frac{1}{2}N(N-1)$ pairs of observations. Stuart (1954) has shown that

$$R^2(V) \sim \frac{N^3}{4\pi}, \tag{18}$$

so that using (15) and (18) in (2) and (4) with $r = 3$, we obtain for the A.R.E. of S_1 compared to V

$$\text{A.R.E.}\,(S_1, V) = (\tfrac{8}{9})^{\frac{1}{2}} = 0{\cdot}87. \tag{19}$$

The loss of A.R.E. involved in reducing the number of comparisons from $\frac{1}{2}N(N-1)$ to $\frac{1}{2}N$ is as little as 13%.

These values of the A.R.E. depend on the assumption of normality, but the calculation of the form of the optimum statistic, S_1, and also of the statistic, S_3, of § 5, does not. For (7) remains true, with a changed numerical factor, for general continuous distributions.

5. An unweighted sign test

The relatively high efficiency of S_1 compared to V leads us to construct, by analogy, a simplified version of Kendall's rank correlation test, which may be defined by

$$Q = \sum_{i<j} h_{ij}, \tag{20}$$

and gives equal weight to all $\frac{1}{2}N(N-1)$ comparisons. Q has the same A.R.E. as V (Stuart, 1954). The analogous sign test, based on $\frac{1}{2}N$ equally weighted independent comparisons, is

$$S_2 = \Sigma h_{ij}, \tag{21}$$

and using (10), we obtain, with all $w_{ij} = 1$,

$$R^2(S_2) = \frac{2}{N\pi}(\Sigma r_{ij})^2. \tag{22}$$

We now require to choose $\frac{1}{2}N$ pairs from the first N integers so that (22) or, equivalently, $\Sigma(j-i) = \Sigma r_{ij}$, takes its maximum value. This occurs whenever every i is chosen from the first $\frac{1}{2}N$ integers and every j from the last $\frac{1}{2}N$ integers. In particular, it occurs when every $r_{ij} = \frac{1}{2}N$ exactly, so that

$$S_2 = \sum_{k=1}^{\frac{1}{2}N} h_{k,\,\frac{1}{2}N+k}, \tag{23}$$

and (22) becomes

$$R^2(S_2) = \frac{N^3}{8\pi}. \tag{24}$$

Using (24) and (18), we obtain

$$\text{A.R.E.}\,(S_2, V) = (\tfrac{1}{2})^{\frac{1}{2}} = 0{\cdot}79, \tag{25}$$

while from (15)

$$\text{A.R.E.}\,(S_2, S_1) = (\tfrac{1}{2})^{\frac{1}{2}} = 0{\cdot}91. \tag{26}$$

Thus the simplified version of Kendall's rank correlation test is 21 % less efficient than Q or V, and 9 % less efficient than the simplified Spearman coefficient S_1. The use of S_2 is equivalent to a test considered by Theil (1950).

6. The best unweighted sign test

However, we can improve on the efficiency of S_2, and in fact get very nearly as high an efficiency as that of S_1, by 'throwing away' some of the $\frac{1}{2}N$ comparisons and retaining equal weights for the others. This was suggested by one of the present authors in the discussion of Foster & Stuart (1954); it leads to an increase in efficiency because, by comparing observations further apart, individual comparisons are made more sensitive.

In (1), let every w_{ij} be either 0 or 1, and let m ($\leqslant \frac{1}{2}N$) be the number of non-zero w_{ij}. For our new statistic S_3 we have, as in (8),

$$E'(S_3) = -\frac{1}{2\sqrt{\pi}} \Sigma w_{ij} r_{ij} \quad (w_{ij} = 0 \text{ or } 1), \tag{27}$$

and from (9)
$$D^2(S_3 \mid \Delta = 0) = \tfrac{1}{4}m, \tag{28}$$

so that (3), (27) and (28) give

$$R^2(S_3) = \frac{1}{m\pi} (\Sigma w_{ij} r_{ij})^2 \quad (w_{ij} = 0 \text{ or } 1). \tag{29}$$

To maximize this by choice of m and r_{ij}, we again work in two stages. For fixed m, (29) will take its largest value when the comparisons given zero weights are based on the middle $(N-2m)$ observations, while every i is chosen from the first m observations and every j is chosen from the last m observations. In particular, this will be so when every $r_{ij} = (N-m)$ exactly, so that

$$S_3 = \sum_{k=1}^{m} h_{k, N-m+k} \tag{30}$$

and (29) becomes
$$R^2(S_3) = \frac{m(N-m)^2}{\pi}. \tag{31}$$

(31) is the largest possible value of $R^2(S_3)$ for fixed m. (S_2 is the special case of S_3 when $m = \frac{1}{2}N$.) We now choose m to maximize (31). Differentiating, we get

$$m = \tfrac{1}{3}N \tag{32}$$

for a maximum, so that finally

$$S_3 = \sum_{k=1}^{\frac{1}{3}N} h_{k, \frac{1}{3}N+k},$$

for which
$$\left.\begin{aligned} E(S_3) &= \tfrac{1}{6}N, \\ V(S_3) &= \tfrac{1}{12}N, \end{aligned}\right\} \tag{33}$$

and from (31),
$$R^2(S_3) = \frac{4N^3}{27\pi}. \tag{34}$$

From (34) and (18), we have

$$\text{A.R.E.}\,(S_3, V) = (\tfrac{8}{27})^{\frac{1}{2}} = 0{\cdot}84, \tag{35}$$

while from (15)
$$\text{A.R.E.}\,(S_3, S_1) = (\tfrac{8}{9})^{\frac{1}{2}} = 0{\cdot}96. \tag{36}$$

Compared with either V or S_1, S_3 has about 5 % higher efficiency than S_2, and in fact its efficiency is 96 % of that of S_1, so that for practical purposes it may be recommended instead of S_1 because it requires no weighting of the comparisons.

7. Comparison of the sign tests

In Table 1, the A.R.E. of the sign tests are tabulated, compared to each other, to the rank correlation tests, and to the best (parametric) test against normal regression, based on the sample regression coefficient b, which has a value of (3) given by

$$R^2(b) \sim \frac{N^3}{12}, \tag{37}$$

as follows immediately from the fact that b is an unbiased estimator of Δ with variance $12/\{N(N^2-1)\}$.

Table 1. *Asymptotic relative efficiencies of sign tests*

Test statistic	Asymptotic relative efficiency		
	Compared to S_1	Compared to rank correlation tests	Compared to best parametric test
$S_1 = \sum_{k=1}^{\frac{1}{2}N} (N-2k+1)\, h_{k,\, N-k+1}$	1·00	0·87	0·86
$S_2 = \sum_{k=1}^{\frac{1}{2}N} h_{k,\, \frac{1}{2}N+k}$	0·91	0·79	0·78
$S_{\frac{1}{2}} = \sum_{k=1}^{\frac{1}{2}N} h_{k,\, \frac{1}{2}N+k}$	0·96	0·84	0·83

From (2), (18) and (37), it follows that the A.R.E. of either rank correlation coefficient compared to b is

$$\text{A.R.E.}\,(V, b) = \left(\frac{3}{\pi}\right)^{\frac{1}{2}} = 0{\cdot}98, \tag{38}$$

and not $3/\pi = 0{\cdot}95$ as given by Stuart (1954).

8. Comparison with A.R.E. of other tests

Apart from the two rank correlation tests already discussed, Stuart (1954) investigated the A.R.E. of three other distribution-free tests for trend in location. Two of these, the rank serial correlation test and the turning point test, were found to have zero values of R as defined by (3); the third, the difference-sign test, was found to have a value of r equal to 1 in (4), as against $r = 3$ for all the tests considered in this paper. It followed that the three tests mentioned all have A.R.E. zero compared to the rank correlation tests (and hence to all the tests discussed here). Noether (1955) gives general results which rigorize these conclusions.

A well-known and simple test which has not, as far as we know, previously been discussed from the point of view of A.R.E. is the median test, due to Brown & Mood (1951). The N (even) observations are divided into two sets of $\frac{1}{2}N$ consecutive observations. The test

statistic is simply the number of observations in the first set which exceed the sample median y_m, and it is therefore defined by

$$B = \sum_{i=1}^{\frac{1}{2}N} b_{im}, \tag{39}$$

where
$$b_{im} = \begin{cases} 1 & \text{if } y_i > y_m, \\ 0 & \text{if } y_i < y_m. \end{cases}$$

The A.R.E. of B is easily obtained. We know that y_i is a normal variate with mean $(\alpha + i\Delta)$ and unit variance. It follows that the sample median y_m is asymptotically a normal variate with mean $(\alpha + \frac{1}{2}(N+1)\Delta)$ and variance of order N^{-1}. Since y_i and y_m are asymptotically independent, $(y_i - y_m)$ is asymptotically normal with mean $\Delta[i - \frac{1}{2}(N+1)]$ and unit variance, so that for $i < \frac{1}{2}(N+1)$

$$E(b_{im}) = \text{prob}\,(y_i > y_m) \sim 1 - G\{\Delta[\tfrac{1}{2}(N+1) - i]\}. \tag{40}$$

Using (6) in (40), we obtain

$$E'(b_{im}) \sim -\frac{1}{\sqrt{(2\pi)}}[\tfrac{1}{2}(N+1) - i], \tag{41}$$

so that from (39) and (41),

$$E'(B) = \sum_{i=1}^{\frac{1}{2}N} E'(b_{im}) \sim -\frac{1}{\sqrt{(2\pi)}} \sum_{i=1}^{\frac{1}{2}N} [\tfrac{1}{2}(N+1) - i] \sim -\frac{N^2}{8\sqrt{(2\pi)}}. \tag{42}$$

Also (Brown & Mood, 1951)
$$D^2(B \mid \Delta = 0) \sim \frac{N}{16}. \tag{43}$$

(42) and (43) give, in (3),

$$R^2(B) \sim \frac{N^3}{8\pi}. \tag{44}$$

Comparison of (44) with (24) shows that B has precisely the same A.R.E. as S_2, and is therefore slightly less efficient than S_1 and S_3. If the observations are available in serial order, S_3 is simpler to compute than B, which involves ranking all the observations to find the median, and then making $\frac{1}{2}N$ comparisons, as against $\frac{1}{2}N$ for S_3. There is therefore no reason to prefer B to S_3 in this case. If, however, the data were available graphically, B would be considerably easier to compute, and this would outweigh the slight loss of efficiency compared to S_3.

9. Comparison of the powers of tests

So far we have compared tests by the A.R.E. in the usual way. Before considering the power of the test S_3 in small samples it is convenient to examine the meaning of the A.R.E. more carefully. If the A.R.E. of a quick test relative to an efficient test is A, then asymptotically A^{-1} as many observations have to be made for the quick test to give the same local power as the efficient test. This is directly relevant if in designing an experiment a choice has to be made between, on the one hand, using an efficient method of analysis and on the other taking more observations and using a quick method of analysis. But it is not directly relevant to the choice of a method of analysis for a given body of data, because it depends in part on r, defined by (4), measuring the rate at which power increases with increasing N. For a given problem r is fixed and so the A.R.E. can be reinterpreted in terms of the power attained at a fixed sample size, but it seems preferable to compare tests directly in terms of power.

Consider a test based on a statistic S normally distributed with mean $E(S \mid \Delta)$ and standard deviation $D(S \mid \Delta)$, where the null hypothesis is $\Delta = 0$. The null hypothesis is rejected at the significance level α if

$$S > E(S \mid 0) + \lambda_\alpha D(S \mid 0), \tag{45}$$

where

$$G(-\lambda_\alpha) = \alpha. \tag{46}$$

The power of the test is $G[p(\Delta)]$, where

$$p(\Delta) = \frac{E(S \mid \Delta) - E(S \mid 0) - \lambda_\alpha D(S \mid 0)}{D(S \mid \Delta)}. \tag{47}$$

Now

$$p'(0) = \left(\frac{\partial p(\Delta)}{\partial \Delta}\right)_{\Delta=0} = \frac{E'(S \mid 0) + \lambda_\alpha D'(S \mid 0)}{D(S \mid 0)}. \tag{48}$$

In all the applications in this paper $D'(S \mid 0) = 0$, so that

$$p'(0) = \frac{E'(S \mid 0)}{D(S \mid 0)} = R(S). \tag{49}$$

Near $\Delta = 0$,

$$p(\Delta) = \Delta R(S) - \lambda_\alpha + O(\Delta^2), \tag{50}$$

and in applications the first two terms give, asymptotically in N, the whole of the power curve. Moreover, $R(S) \sim RN^{-\frac{1}{2}r}$ as $N \to \infty$ and comparable tests of a given hypothesis will have the same r; hence we usually need to consider just R. We call $p(\Delta)$ the *power deviate* and $p'(0)$ the *power derivative*. Asymptotically the graph of $p(\Delta)$ against Δ is linear, and tests at different significance levels are given by parallel straight lines. Or to put the same fact another way, the power curves are asymptotically linear when plotted on arithmetical probability paper.

Now consider the small sample theory with S possibly not normally distributed. Then if the power curves are plotted on probability paper they can be expected to form an approximately parallel set of curves approaching a set of parallel lines as the sample size increases and the distribution of S tends to normality. This is of course only a method of presenting the results of power calculations, but we shall find it very convenient both in assessing the small-sample behaviour and in comparing different tests.

Consider now two tests for which the asymptotic values of $R(S)$ are R_1, R_2. Then asymptotically in N the power curves for a given α are two lines on probability paper, the ratio of their slopes being R_1/R_2 independent of α; both lines intersect the probability axis at α.

A first consequence is that there is no simple general relation between the difference in the power of the two tests and the ratio R_1/R_2. If $R_1 \neq R_2$ we can, by taking α sufficiently small, make the difference in power between the two tests arbitrarily near unity. In practice we are probably only interested in $0{\cdot}20 \geqslant \alpha \geqslant 0{\cdot}001$, but the general conclusion remains that the difference in power between a quick test and an efficient test will be greatest for small α. Table 2 expresses this quantitatively; it shows for given R_1/R_2 the powers of the two tests at the point at which the difference in powers is greatest. The values in Table 2 are independent of N, but the values of Δ at which these powers are attained do depend on N. This is the restriction on the alternative hypothesis referred to in § 3. Thus if $R_1/R_2 = 0{\cdot}7$ and $\alpha = 0{\cdot}05$, the difference in power is greatest for the value of Δ at which the power of the efficient test is 77 % and of the quick test 51 %.

Now consider the power of S_3 in small samples. Two methods can be used. The first is to take the expansion (50) to higher powers of Δ and to introduce a correction for the non-normality of S based on an Edgeworth expansion. This may be shown to give good results even for very small N, and is a general method which could be used where direct numerical calculation is difficult. However, for S_3 it is much easier to calculate the power directly from the National Bureau of Standards tables of the binomial distribution (1950).

Table 2. *Asymptotic theory. Powers (per cent) of quick and efficient tests at points at which difference in power is greatest**

R_1/R_2 \ α	0·10	0·05	0·01	0·001
0·9	67, 73	63, 71	49, 60	54, 67
0·8	61, 74	56, 72	49, 71	43, 72
0·7	59, 80	51, 77	42, 77	39, 83
0·6	54, 84	47, 84	39, 86	29, 87
0·5	48, 88	41, 89	30, 90	20, 93
0·3	35, 96	27, 96	14, 97	7, 99

The power was computed in this way for $N = 15\,(15)\,135$, the significance level being the largest value $\leqslant 0{\cdot}05$. Under the null hypothesis the test statistic is distributed as $(\frac{1}{2}+\frac{1}{2})^{\frac{1}{2}N}$ and under the alternative hypothesis as $(p+q)^{\frac{1}{2}N}$, where

$$p = G\left(-\frac{\sqrt{2}N\Delta}{3}\right). \tag{51}$$

The power corresponding to given p, $\frac{1}{2}N$ can be read off directly from the tables and (51) solved for Δ. The results are given in Table 3. For comparative purposes the exact power of the parametric test based on the regression coefficient, b, has been computed for the same values of N and Δ. When the standard deviation about the regression line is known, the power is exactly $G[p(\Delta)]$, where

$$p(\Delta) = \{\tfrac{1}{12}N(N^2-1)\}^{\frac{1}{2}}\Delta - \lambda_\alpha. \tag{52}$$

To avoid rewriting the values of Δ in Table 4 the rows of both tables have been lettered, and each entry in Table 4 relates to the value of Δ shown above the corresponding entry in Table 3.

Asymptotically, the ratio of the R values of the two tests is, by (34) and (37), $4/(3\sqrt{\pi}) = 0{\cdot}75$; the interpretation of this in terms of power can be obtained from Table 2. The full curve in Fig. 1 and the full curves in Fig. 2 for $k = 0$ show the power curves for $N = 15$, 30, and the dotted lines are the corresponding asymptotic power curves. The small-sample power is lower than the value given by the asymptotic theory, the difference being quite appreciable in the region of 80–90 % power. The power curves of the most efficient test are exactly linear and differ from their asymptotic form only because of the very small difference between $\{N(N^2-1)\}^{\frac{1}{2}}$ and $N^{\frac{1}{2}}$. Hence the test S_3 is less efficient relative to b than the asymptotic

* These values were obtained graphically by drawing on probability paper lines whose ratio of slopes is R_1/R_2 and reading off the maximum difference in probability between them. The differences in power are determined accurately, but it is rather difficult to find the precise point of maximum difference. The values in Table 2 involve R_1, R_2 only through their ratio R_1/R_2.

Table 3. *Exact power of S_3 test against normal regression alternatives*

Values of the standardized regression coefficient, Δ, are given in parentheses, and the corresponding power appears immediately below.

Sample size (N) ...	15	30	45	60	75	90	105	120	135
Significance level α ...	0·031	0·011	0·018	0·021	0·022	0·049	0·045	0·040	0·036
a	(0·0035) 0·035	(0·0018) 0·013	(0·0012) 0·021	(0·0009) 0·026	(0·0007) 0·027	(0·0006) 0·062	(0·0005) 0·057	(0·0004) 0·053	(0·0004) 0·048
b	(0·0178) 0·050	(0·0089) 0·023	(0·0059) 0·042	(0·0044) 0·055	(0·0036) 0·064	(0·0030) 0·135	(0·0025) 0·134	(0·0022) 0·133	(0·0020) 0·130
c	(0·0358) 0·078	(0·0179) 0·046	(0·0119) 0·091	(0·0090) 0·126	(0·0072) 0·154	(0·0060) 0·291	(0·0051) 0·306	(0·0045) 0·317	(0·0040) 0·327
d	(0·0545) 0·116	(0·0272) 0·086	(0·0182) 0·173	(0·0136) 0·245	(0·0109) 0·306	(0·0091) 0·508	(0·0078) 0·542	(0·0068) 0·572	(0·0060) 0·598
e	(0·0742) 0·168	(0·0371) 0·149	(0·0247) 0·297	(0·0185) 0·416	(0·0148) 0·512	(0·0124) 0·730	(0·0106) 0·773	(0·0093) 0·807	(0·0082) 0·836
f	(0·0954) 0·237	(0·0477) 0·244	(0·0318) 0·461	(0·0238) 0·617	(0·0191) 0·727	(0·0159) 0·894	(0·0136) 0·924	(0·0119) 0·946	(0·0106) 0·961
g	(0·1190) 0·328	(0·0595) 0·376	(0·0397) 0·648	(0·0298) 0·804	(0·0238) 0·891	(0·0198) 0·974	(0·0170) 0·986	(0·0149) 0·992	(0·0132) 0·996
h	(0·1466) 0·444	(0·0733) 0·544	(0·0489) 0·823	(0·0366) 0·933	(0·0293) 0·975	(0·0244) 0·997	(0·0209) 0·999	(0·0183) 1·000	(0·0163) 1·000
i	(0·1812) 0·590	(0·0906) 0·736	(0·0604) 0·944	(0·0453) 0·989	(0·0362) 0·998	(0·0302) 1·000	(0·0259) 1·000	(0·0227) 1·000	(0·0201) 1·000
j	(0·2326) 0·774	(0·1163) 0·914	(0·0775) 0·995	(0·0582) 1·000	(0·0465) 1·000	(0·0388) 1·000	(0·0332) 1·000	(0·0291) 1·000	(0·0258) 1·000

Table 4. *Exact power of b test against normal regression alternatives*

Values of Δ are given in parentheses above the corresponding entry in Table 3.

Sample size (N) ...	15	30	45	60	75	90	105	120	135
Significance level α ...	0·031	0·011	0·018	0·021	0·022	0·049	0·045	0·040	0·036
a	0·036	0·013	0·023	0·027	0·030	0·066	0·062	0·057	0·054
b	0·059	0·030	0·056	0·074	0·088	0·179	0·182	0·183	0·183
c	0·103	0·074	0·143	0·201	0·249	0·429	0·457	0·481	0·502
d	0·171	0·157	0·300	0·416	0·509	0·722	0·764	0·799	0·827
e	0·267	0·294	0·519	0·673	0·775	0·919	0·945	0·962	0·973
f	0·395	0·485	0·746	0·877	0·941	0·988	0·994	0·997	0·999
g	0·551	0·699	0·911	0·975	0·993	0·999	1·000	1·000	1·000
h	0·772	0·880	0·984	0·998	1·000	1·000	1·000	1·000	1·000
i	0·879	0·977	0·999	1·000	1·000	1·000	1·000	1·000	1·000
j	0·979	0·999	1·000	1·000	1·000	1·000	1·000	1·000	1·000

theory suggests. The difference is greater for smaller α. The corresponding graphs to Figs. 1 and 2 for higher N show that for α in the range 0·01–0·05, the asymptotic theory applies well for $N \geqslant 60$.

The next thing is to investigate whether the form of S_3, involving the rejection of the middle third of the set of observations, can profitably be modified in small samples. Suppose that $(\frac{1}{3}N - 2k)$ observations are rejected so that the number of comparisons is $(\frac{1}{3}N + k)$; the exact power function can be worked out from the binomial tables as before, but an immediate comparison is not possible because the significance levels for different values of α cannot be made equal, except by the artificial device of randomized tests. However, if the curves are plotted on probability paper they are almost parallel for different α and an

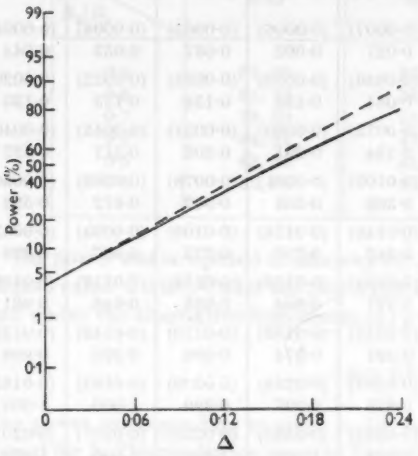


Fig. 1. Power of S_3 for $N = 15$. —— Exact power. ---- Value from asymptotic theory.

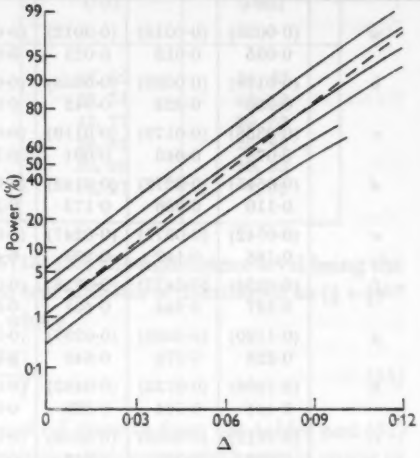


Fig. 2. Power of S_3 for $N = 30$. —— Exact power. ---- Value from asymptotic theory. Full curves are in descending order $k = 0, 1, 2$. 0, 1, 2, the second set having lower values of α than the first. Broken curve is $k = 0$.

increase in power with change in k would be shown by decreasing curvature. As would be expected, a negative k leads to a loss of power. Fig. 2 shows for $N = 30$ the curves for $k = 0, 1, 2$. There is a tendency for the curvature to increase as α decreases, but there does not seem to be any systematic change with k. Therefore, although an increase in k increases the number of available significance levels, it does not appreciably increase power. Hence there seems to be little value in modifying the $\frac{1}{3}$ rule in small samples; similar calculations for $N = 15$ confirm this.

We have not made the corresponding investigations for S_1.

10. Examples of use of the sign tests against trend in location

To illustrate the S_1 and S_3 tests, we use the figures of annual rainfall at Oxford for the years 1858–1952, quoted by Foster & Stuart (1954, Table 9).

For S_1 we compare the kth observation with the $(N - k + 1)$th, scoring 1 when the former is the larger and 0 when it is the smaller. The unit scores are then weighted by the distance

apart of the observations compared, i.e. by $(N-2k+1)$. In this case $N = 95$ and is odd, so that we must ignore the middle observation and proceed with $N = 94$. The unit scores are those with weights as follows:

89, 83, 79, 77, 75, 71, 65, 59, 57, 55, 51, 49, 47, 45, 33, 31, 27, 15, 3.

The value of S_1 is the sum of these weights, 1011. From (16), with $N = 94$, we have

$$E(S_1) = 1104{\cdot}5, \quad D^2(S_1) = 34603{\cdot}75, \quad D(S_1) = 186{\cdot}0.$$

The observed value of S_1 thus represents a deviation from expectation of almost exactly one-half its standard error and is therefore in good agreement with the null hypothesis of zero trend.

If, alternatively, we were to use the simpler S_3 test we compare the kth observation with the $(\frac{2}{3}N+k)$th, scoring 1 or 0 as before, but no weighting is necessary. Since $N = 95$ and is not a multiple of three, we retain the extra observations (in accordance with our findings at the end of § 9 above) and compare each of the first thirty-two observations with the corresponding observation in the last thirty-two. For our sequence of scores we obtain

1 0 1 0 1 0 0 1 1 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 0 0 0 1 0 0 0,

the total score S_3 being 14. This clearly agrees well with the expected value of $\frac{1}{2} \times 32 = 16$. The standard error of S_3 is, from (33), $\sqrt{(\frac{1}{4} \times 32)} = 2{\cdot}83$, so that the deviation from expectation, corrected for continuity, is just over one-half of a standard error.

11. Sign tests for trend in dispersion

We consider now tests for a trend not in location but in the dispersion about a fixed location. For example, in a regression problem we may want to test quickly whether the scatter about the regression curve increases as the independent variable increases.

Divide the series $x_1, \ldots, x_N$ into sets $x_1, \ldots, x_k; x_{k+1}, \ldots, x_{2k}; \ldots$, rejecting a few observations in the centre of the original series if N is not exactly divisible by k. The best choice of k is discussed below. For each set of k observations find the range, w, thus getting a series of ranges $w_1, \ldots, w_r$, where r is the integral part of N/k. The ranges are then tested for trend by one or other of the tests S_1 and S_3.

If the null hypothesis is that $x_1, \ldots, x_N$ are independently distributed with constant dispersion about a regression line, $w_1, \ldots, w_r$ are independent and identically distributed. If the regression is not linear the w's will be approximately identically distributed unless the trend within sets of observations varies appreciably.

In the next section, a valid test is obtained for any k, and the best value of k for detecting certain special forms of trend is found for large samples. The behaviour for small samples has not been investigated. The following provisional rules are suggested:

if	$N \geqslant 90$	take	$k = 5$,
if	$90 > N \geqslant 64$	take	$k = 4$,
if	$64 > N \geqslant 48$	take	$k = 3$,
if	$48 > N$	take	$k = 2$.

Except when N is very large it is probably advisable to use the weighted, rather than the unweighted, sign test.

12. Choice of k for dispersion tests

To investigate the theory of the test for trend in dispersion we take a special form for the null and alternative hypotheses. Suppose that $x_1, \ldots, x_N$ are independently normally distributed with constant mean and with standard deviations $\sigma(1), \ldots, \sigma(N)$, where $\sigma(n)$ varies at most slowly with n. Then the ranges $w_1, \ldots, w_r$ defined in § 11 have very nearly the distribution of ranges of k observations drawn from normal populations of standard deviations $\sigma_1 = \sigma(\frac{1}{2}k)$, $\sigma_2 = \sigma(\frac{3}{2}k)$, ... Patnaik (1950) has shown that a range of k observations can be represented to a close approximation as a multiple of a χ-variate with suitably chosen degrees of freedom, ν_k. Therefore $w_i^2\sigma_j^2/(w_j^2\sigma_i^2)$ is approximately an F variate with (ν_k, ν_k) degrees of freedom.

Hence
$$\text{prob}\,(w_i/w_j > 1) \simeq \int_0^{(\sigma_i/\sigma_j)^2} \frac{\Gamma(\nu_k)}{[\Gamma(\frac{1}{2}\nu_k)]^2} \frac{x^{\frac{1}{2}\nu_k - 1}}{(1+x)^{\nu_k}}\,dx$$
$$\simeq \frac{1}{2} + \left\{\left(\frac{\sigma_i}{\sigma_j}\right)^2 - 1\right\} A_k$$

where
$$A_k = \frac{\Gamma(\nu_k)}{[\Gamma(\frac{1}{2}\nu_k)]^2}\,(\tfrac{1}{2})^{\nu_k},$$

provided that $(\sigma_i/\sigma_j)^2 - 1$ is small.

If we assume that the trend in standard deviation is such that $\sigma(n) = \sigma_0 e^{\gamma n} \sim \sigma_0(1+\gamma n)$, where $\gamma N \ll 1$, so that γ is the fractional increase in standard deviation per observation, we have
$$(\sigma_i/\sigma_j)^2 - 1 \sim 2k\gamma(i-j)$$

and
$$\text{prob}\,(w_i/w_j > 1) \simeq \tfrac{1}{2} + 2k\gamma(i-j)\,A_k.$$

Consider first the application to the ranges of the unweighted sign test, S_3. From the $r \simeq N/k$ ranges we make approximately $\frac{1}{3}N/k$ independent comparisons in each of which $i - j \simeq \frac{2}{3}r$. Therefore if S is the total score, its mean and standard deviation are given by

$$E(S \mid \gamma) \simeq \frac{1}{3}\frac{N}{k}\,4k\gamma\,\frac{2rA_k}{3} = \frac{8N^2A_k\gamma}{9k},$$

$$D(S \mid \gamma) = \left(\frac{N}{3k}\right)^{\frac{1}{2}} + O(\gamma^2).$$

Therefore the power derivative, $p_3'(0)$, of the test is

$$p_3'(0) = \frac{E'(S \mid 0)}{D(S \mid 0)} \simeq \frac{8\sqrt{3}\,N^{\frac{1}{2}}A_k}{9\sqrt{k}}. \tag{53}$$

An exactly analogous calculation for the weighted sign test S_1 gives

$$p_1'(0) \simeq \frac{2\sqrt{6}\,N^{\frac{1}{2}}A_k}{3\sqrt{k}}. \tag{54}$$

Thus in both cases the asymptotically best value of k is the one that maximizes $A_k/\sqrt{k}$. From Patnaik's table of ν_k the values in Table 5 have been computed.

Now the number of ranges is N/k and the number of comparisons is one-half or one-third of this and is therefore small even when N is, by usual standards, quite large. Therefore it is advisable to use smaller values of k than the theoretical optimum in large samples. In the absence of an investigation of the small-sample properties of the test, the rule of § 11

Table 5. *Determination of efficiencies of different set sizes, k, for testing trend in dispersion*

k	$A_k/\sqrt{k}$	k	$A_k/\sqrt{k}$
2	0·112	6	0·167
3	0·141	7	0·169
4	0·158	8	0·170
5	0·164	9	0·169

is suggested. This is based on the considerations that there is little gain in taking $k > 5$ and that it is advisable, whenever possible, to have at least sixteen ranges.

If we substitute, in (53) and (54), the value $A_k/\sqrt{k} \simeq 0{\cdot}16$, we have

$$\left.\begin{aligned} p_3'(0) &\simeq 0{\cdot}246N^{\frac{1}{2}}, \\ p_1'(0) &\simeq 0{\cdot}261N^{\frac{1}{2}}. \end{aligned}\right\} \tag{55}$$

It remains to compare (55) with the corresponding quantity for the maximum-likelihood test of the corresponding parametric hypotheses.

13. A.R.E. of dispersion tests

For simplicity assume that $x_1, \ldots, x_N$ are normally and independently distributed with zero mean and that the standard deviation of x_n is $\sigma_0 e^{\gamma n}$, where γN is small. The log likelihood is

$$L = -\tfrac{1}{2}N\log(2\pi) - N\log\sigma_0 - \gamma\sum_1^N n - \frac{1}{2\sigma_n^2}\Sigma x_n^2 e^{-2\gamma n}.$$

If we differentiate and take expectations, retaining only the terms independent of γ, and letting N tend to infinity, we get

$$E\left(\frac{\partial^2 L}{\partial\sigma_0^2}\right) \sim -\frac{2N}{\sigma_0^2}, \quad E\left(\frac{\partial^2 L}{\partial\sigma_0\partial\gamma}\right) \sim -\frac{N^2}{\sigma_0}, \quad E\left(\frac{\partial^2 L}{\partial\gamma^2}\right) \sim -\tfrac{2}{3}N^3. \tag{56}$$

The large-sample variance of $\hat{\gamma}$, the maximum-likelihood estimate of γ, is given by inverting the Hessian matrix with elements (56). We get when γ is small and N is large

$$V(\hat{\gamma}) \sim 6/N^3. \tag{57}$$

Thus the power derivative of the test based on the maximum-likelihood estimate is

$$p_m'(0) = \frac{N^{\frac{1}{2}}}{\sqrt{6}} = 0{\cdot}408N^{\frac{1}{2}}. \tag{58}$$

(58) still applies if the mean in unknown or if a linear trend in mean has to be estimated.

From the formulae (55) and (58), and the fact that $p_x'(0) = R(x)$, it follows, on using (2) with $r = 3$, that the A.R.E.'s of the tests S_3, S_1 compared with the maximum-likelihood test are about 71 and 74 % respectively.

A test entirely analogous to the above tests can be found by calculating the variances within each set of k instead of the range. This is slightly more efficient in the parametric case but much of the simplicity of the test is lost and the increase in power may be shown to be trivial.

14. Examples of the use of sign tests against trend in dispersion

We again use for illustrative purposes the rainfall data quoted by Foster & Stuart (1954, Table 9). Using the provisional rule given in § 11 above, we take ranges of sets of five observations. Since $N = 95$, this gives us exactly nineteen sets, no rejection of observations being necessary. The nineteen ranges are:

9·64, 12·30, 12·01, 11·45, 5·43, 13·05, 9·86, 10·89, 6·95, 15·03,
11·34, 6·63, 12·19, 8·55, 4·80, 11·00, 7·76, 7·03, 10·98.

If we apply the test S_1 we drop the middle value and take the signs of 10·98–9·64, 7·03–12·30, ... down to 11·34–6·95, thus obtaining

score:	0	1	1	1	1	1	0	1	0
weight:	17	15	13	11	9	7	5	3	1

The total score is therefore 58, and from (16) with $N = 18$ the mean score is 40·5 and the variance is 242·25, so that the standard error is 15·6. The deviation from expectation is about 1·12 standard errors, and so the two-sided normal significance level is about 27 %. The exact significance level is, by enumeration, $73/256 \simeq 28\frac{1}{2}$ %.

If we use the test S_3 we reject the middle five of the nineteen ranges and take the signs of 12·19–9·64, etc. This gives scores

0 1 1 1 0 1 0

There is clearly good agreement with an equal probability for zeros and ones; significance would be tested in the binomial distribution $(\frac{1}{2} + \frac{1}{2})^7$. The test S_3 is not to be recommended in the present instance because with only seven comparisons the loss of sensitivity compared to the S_1 test would be considerable.

Thus although there is a slight indication that the dispersion decreases with time, both tests suggest that this could easily be a sampling fluctuation.

15. Sequential tests

Finally, we point out the possibility of constructing sequential tests for trend related to the tests considered above. While this paper was in preparation an abstract (Noether, 1954) appeared describing briefly a test rather similar to the one we had developed. Hence a full discussion will not be attempted here. However, some calculations in a special case suggest that the average sample size under the null and alternative hypotheses are, for the sequential sign tests, only a little greater than the corresponding parametric fixed sample size.

Sequential sign tests for trend are only likely to be of practical value under rather exceptional circumstances. For they require that observations are sufficiently easy to obtain for it to be worth while to use inefficient methods of analysis, and yet sufficiently difficult to obtain for the saving from the use of a sequential method to be important. A possible application is to the marking of a large number of examination scripts. If they are marked in alphabetical order it may be useful to test, as the marking proceeds, for a trend in the marks, which would indicate a changing standard of marking. A sequential method is appropriate and yet elaborate calculations would be out of place.

16. General comments

The calculation of the efficiency of the above tests and the determination of optimum weightings, etc., has been based on a particular type of alternative hypothesis. It is clear in a general way that the tests will remain effective for detecting monotone trends. Positive serial correlation among the observations would increase the chance of a significant answer even in the absence of a trend.

The occurrence of ties has been ignored in the above work. A small number of ties can be dealt with by counting one-half a comparison in each direction, i.e. if $y_i = y_j$ we calculate as if one-half a comparison has $y_i > y_j$ and one-half has $y_i < y_j$. If a substantial proportion of the comparisons are ties a special investigation is necessary or the comparisons should be randomized.

Estimates for the trend could be constructed from the test statistics S_1, S_3. It is very doubtful if such estimates would be of value; in any case, in much work with quick tests, if the trend is shown to be significant it can be estimated graphically.

References

Brown, G. W. & Mood, A. M. (1951). On median tests for linear hypotheses. *Proceedings of the Second Berkeley Symposium*, pp. 159–66. University of California Press.

Foster, F. G. & Stuart, A. (1954). Distribution-free tests in time-series based on the breaking of records. *J. R. Statist. Soc.* Series B (Methodological), **16**, 1–22.

National Bureau of Standards (1950). *Tables of the Binomial Probability Distribution*, Applied Mathematics Series, no. 6. Washington.

Noether, Gottfried, E. (1954). Abstract of 'A sequential test of randomness against linear trend'. *Ann. Math. Statist.* **25**, 176.

Noether, Gottfried E. (1955). 'The asymptotic relative efficiencies of tests of hypotheses'. *Ann. Math. Statist.* (to be published).

Patnaik, P. B. (1950). The use of mean range as an estimator of variance in statistical tests. *Biometrika*, **37**, 78–87.

Pitman, E. J. G. (1948). Lecture notes on 'Non-parametric statistical inference', University of North Carolina Institute of Statistics (mimeographed).

Stuart, Alan (1954). The asymptotic relative efficiencies of distribution-free tests of randomness against normal alternatives. *J. Amer. Statist. Ass.* **49**, 147–57.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Indag. math.* **12**, 85–91, 173–7.

THE VARIANCE OF THE MAXIMUM OF PARTIAL SUMS OF A FINITE NUMBER OF INDEPENDENT NORMAL VARIATES

By A. A. ANIS

Cairo University

1. The problem

Consider n independent standard normal variates $X_1, X_2, \ldots, X_n$ and their partial sums

$$S_r = X_1 + \ldots + X_r \quad (r = 1, 2, \ldots, n).$$

Let $$U_n = \operatorname*{Max}_r \{S_r\}$$

denote the maximum of these partial sums.

In a paper by Anis & Lloyd (1953) the expectation of U_n was studied and it was shown that

$$\mathscr{E}(U_n) = (2\pi)^{-\frac{1}{2}} \sum_{1}^{\frac{1}{2}n-1} r^{-\frac{1}{2}}.$$

In the present paper the second moment about zero and hence the variance of U_n is obtained (see equation (7·1)).

We shall always use the symbol $\phi(x)$ to denote the probability density function of the standard normal variate, i.e.

$$\phi(x) = (2\pi)^{-\frac{1}{2}} \exp(-\tfrac{1}{2}x^2),$$

and $F_n(x), f_n(x)$ to denote respectively the distribution function and the probability density function of U_n:

$$F_n(x) = Pr(U_n \leqslant x), \quad f_n(x) = \frac{d}{dx} F_n(x).$$

We have

$$F_n(y) = \int_K (n)! \int \prod_1^n \phi(x_i)\, dx_i \quad (-\infty \leqslant y \leqslant \infty), \tag{1·1}$$

where the region of integration K is defined by

$$K: \quad \sum_1^r x_i \leqslant y \quad (r = 1, 2, \ldots, n).$$

It may be deduced that

$$F_n(x) = \int_0^\infty F_{n-1}(t)\, \phi(x-t)\, dt, \tag{1·2}$$

and that

$$f_n(x) = F_{n-1}(0)\, \phi(x) + \int_0^\infty f_{n-1}(t)\, \phi(x-t)\, dt. \tag{1·3}$$

2. Three lemmas on the $F_r(0)$

At this stage we state three results relating to the $F_r(0)$ which we shall need in the sequel.

Lemma 1.

$$\sum_{s=0}^{n} F_s(0)\, F_{n-s}(0) = 1. \tag{2·1}$$

This was proved in Anis & Lloyd (1953), and is repeated here merely for completeness.

LEMMA 2. $$F_r(0) = (2r)!/2^{2r}(r!)^2. \tag{2·2}$$

To prove this we define a generating function

$$M(\lambda) = \sum_{i=0}^{n} \lambda^i F_i(0).$$

Then, using Lemma 1, it is readily seen that

$$M^2(\lambda) = (1-\lambda)^{-1}. \tag{2·3}$$

Picking out the appropriate coefficient from $M(\lambda) = (1-\lambda)^{-\frac{1}{2}}$ gives the required result.

LEMMA 3. $$\sum_{r=0}^{n} rF_r(0)\,F_{n-r}(0) = \tfrac{1}{2}n.$$

This follows from Lemma 2 on differentiating (2·3) and equating coefficients of λ^{n-1}.

3. THE SECOND MOMENT OF U_n AS A LINEAR COMPOUND OF THE $F_r(0)$

The second moment $\mu_2(n)$ of U_n is given by

$$\mu_2(n) = \int_{-\infty}^{\infty} x^2 f_n(x)\,dx,$$

and, using the reduction formula (1·3), this becomes

$$\mu_2(n) = F_{n-1}(0) + \int_{-\infty}^{\infty}\int_{0}^{\infty} x^2\phi(x-t) f_{n-1}(t)\,dt.$$

The double integral can be integrated once, with respect to x. Using well-known properties of $\phi(x)$, and remembering that F_{n-1} is the integral function of f_{n-1}, we obtain

$$\mu_2(n) = 1 + \int_0^{\infty} t^2 f_{n-1}(t)\,dt. \tag{3·1}$$

We now use the reduction formula (1·3) a second time, resulting in

$$\mu_2(n) = 1 + F_{n-2}(0)\int_0^{\infty} t^2\phi(t)\,dt + \int_0^{\infty}\int_0^{\infty} t^2\phi(x-t) f_{n-2}(t)\,dt.$$

The last integral may be reduced in the same way. Continuing, we find

$$\mu_2(n) = 1 + \sum_{r=1}^{n-1} g_r F_{n-r-1}(0), \tag{3·2}$$

where
$$g_r = \int_0^{\infty}(r)\int_0^{\infty} y_1^2\phi(y_1-y_2)\,\phi(y_2-y_3)\ldots\phi(y_{r-1}-y_r)\,\phi(y_r)\prod_1^r dy_i. \tag{3·3}$$

4. THE COEFFICIENTS g_r OF THE LINEAR COMPOUND

We now seek to evaluate the g_r. As a first step we make the transformation

$$\left.\begin{aligned} x_i &= y_i - y_{i+1} \quad (i = 1, 2, \ldots, r;\ y_{r+1} = 0), \\ \text{or} \qquad y_s &= \sum_s^r x_i \qquad (s = 1, 2, \ldots, r). \end{aligned}\right\} \tag{4·1}$$

We then have $$g_r = \int_R (r) \int (x_1 + x_2 + \ldots + x_r)^2 \prod_1^r \phi(x_i)\, dx_i, \tag{4·2}$$

where the region of integration is

$$R: \quad \sum_s^r x_i \geqslant 0 \quad (s = 1, 2, \ldots, r).$$

Hence $$g_r = L_r + H_r,$$

where $$L_r = \int_R (r) \int \left(\sum_1^r x_i^2 \right) \prod_1^r \phi(x_i)\, dx_i, \quad H_r = \int_R (r) \int \left(\sum\sum_{i \neq j} x_i x_j \right) \prod_1^r \phi(x_i)\, dx_i. \tag{4·3}$$

We now consider these two integrals separately.

5. Evaluation of L_r

The integral L_r of (4·3) is readily evaluated, as follows. Since the integrand is spherically symmetrical the value of the integral is proportional to the magnitude of the r-dimensional solid angle defined by the region of integration R; and this in turn is proportional to the $F_r(0)$, since by (1·1) we have

$$F_r(0) = \int_R (r) \int \prod_1^r \phi(x_i)\, dx_i. \tag{5·1}$$

Let us write $F_r^0(0)$ to denote the integral of the integrand of $F_r(0)$ taken over the whole r-space, and let L_r^0 be similarly related to L_r; then the proportionality of our functions $F_r(0)$ and L_r to the solid angle R enable us to write

$$F_r(0) : F_r^0(0) = L_r : L_r^0$$

whence $$L_r = L_r^0 F_r(0), \tag{5·2}$$

since $$F_r^0(0) = 1.$$

The value of L_r^0 is easily obtained by considering the standard integral

$$(2\pi)^{-\frac{1}{2}r} \int_{-\infty}^{\infty} (r) \int_{-\infty}^{\infty} \exp\left(-\tfrac{1}{2} k \sum_1^r x_i^2 \right) \prod_1^r dx_i = k^{-\frac{1}{2}r}.$$

If we differentiate this expression with respect to k and then put $k = 1$ we find

$$L_r^0 = r.$$

Hence, from (5·2), $$L_r = r F_r(0). \tag{5·3}$$

The $F_r(0)$ are, of course, known explicitly.

6. Evaluation of H_r

We now consider the integral H_r of (4·3). We first transform back to the original variables y_i by the transformation (4·1). In this process the cross-product term transforms as follows:

$$\tfrac{1}{2} \sum\sum_{i \neq j} x_i x_j = \sum_1^{r-1} \left\{ x_i \sum_i^r x_j \right\} = \sum_1^{r-1} (y_i - y_{i+1}) y_{i+1}$$

$$= \sum_{s=1}^{r-1} \left\{ (2y_s - y_{s-1} - y_{s+1}) \sum_{s+1}^r y_i \right\},$$

provided we interpret y_0 and y_{r+1} conventionally as

$$y_0 = y_1, \quad y_{r+1} = 0. \tag{6·1}$$

Expressing H_r in terms of the y's, and retaining these conventions, we then have

$$\tfrac{1}{2}H_r = \sum_{s=1}^{r-1} K_s(r), \tag{6·2}$$

where
$$K_s(r) = \int_0^\infty (r) \int_0^\infty (2y_s - y_{s-1} - y_{s+1}) \sum_{s+1}^{r} y_i \prod_{1}^{r} \phi(y_i - y_{i+1})\, dy_i. \tag{6·3}$$

The reason for writing $K_s(r)$ in this form is that it enables us to perform one of the integrations at once, since

$$\int_0^\infty (2y_s - y_{s-1} - y_{s+1})\, \phi(y_{s-1} - y_s)\, \phi(y_s - y_{s+1})\, dy_s = \phi(y_{s-1})\, \phi(y_{s+1}), \tag{6·4}$$

as is immediately seen on using the explicit form of the functions ϕ.

Equation (6·2) then becomes

$$K_s(r) = \int_0^\infty (s-1) \int_0^\infty \prod_{1}^{s-2} \phi(y_i - y_{i+1})\, \phi(y_{s-1})\, dy_1 \dots dy_{s-1}$$
$$\times \int_0^\infty (r-s) \int_0^\infty \left(\sum_{s+1}^{r} y_i \right) \phi(y_{s+1}) \prod_{s+1}^{r} \phi(y_i - y_{i+1})\, dy_{s+1} \dots dy_r. \tag{6·5}$$

Now in this expression the $(s-1)$-fold integral is equal to $F_{s-1}(0)$, as may be seen on applying the transformation (4·1) to (5·1).

The other factor in the last expression, an $(r-s)$-fold integral, we call G_{r-s}. Thus

$$G_k = \int_0^\infty (k) \int_0^\infty \left(\sum_{1}^{k} y_i \right) \phi(y_1)\, \phi(y_1 - y_2) \dots \phi(y_{k-1} - y_k)\, \phi(y_k)\, dy_1 \dots dy_k. \tag{6·6}$$

We now proceed to show that this is expressible in terms of an integral previously evaluated by Anis & Lloyd (1953):

$$E_s = \int_0^\infty (s) \int_0^\infty \phi(y_1)\, \phi(y_1 - y_2) \dots \phi(y_{s-1} - y_s)\, \phi(y_s)\, dy_1 \dots dy_s = (2\pi)^{-\frac{1}{2}} (s+1)^{-\frac{1}{2}}. \tag{6·7}$$

We use the identity
$$\sum_{1}^{k} y_r = \sum_{1}^{k} \tfrac{1}{2} r(k+1-r)\, (2y_r - y_{r-1} - y_{r+1}), \tag{6·8}$$

where, by convention, we take $y_0 = y_{k+1} = 0$.

Equation (6·6) becomes

$$G_k = \sum_{r=1}^{k} \tfrac{1}{2} r(k-r+1) \int_0^\infty (k) \int_0^\infty (2y_r - y_{r-1} - y_{r+1})\, \phi(y_1) \prod_{1}^{k-1} \phi(y_i - y_{i+1})\, \phi(y_k)\, dy_1 \dots dy_k. \tag{6·9}$$

We now use (6·3) to carry out one of the integrations, thus reducing the k-fold integral to

$$\int_0^\infty (k-1) \int_0^\infty \phi(y_1)\, \phi(y_1 - y_2) \dots \phi(y_{r-2} - y_{r-1})\, \phi(y_{r-1})\, \phi(y_{r+1})\, \phi(y_{r+1} - y_{r+2})$$
$$\dots \phi(y_{k-1} - y_k)\, \phi(y_k)\, dy_1 \dots dy_{r-1} dy_{r+1} \dots dy_k$$
$$= E_{r-1} E_{k-r}. \tag{6·10}$$

Gathering the results of this section together, we have from (6·2), (6·5), (6·9) and (6·10),

$$H_r = 2 \sum_{s=1}^{r-1} K_s(r), \quad K_s(r) = F_{s-1}(0)\, G_{r-s}, \quad G_k = \sum_{r=1}^{k} \tfrac{1}{2} r(k+1-r) E_{r-1} E_{k-r}, \quad E_r = (2\pi)^{-\frac{1}{2}} (r+1)^{-\frac{1}{2}}. \tag{6·11}$$

7. Conclusion

We have found (3·2) that the second moment of the maximum of the partial sums $X_1, X_1+X_2, \ldots, X_1+\ldots+X_n$ is

$$\mu_2'(n) = 1+\sum_{r=1}^{n-1} g_r F_{n-r-1}(0),$$

where

$$g_r = rF_r(0)+2\sum_{s=1}^{n-1} F_{s-1}(0)\,G_{r-s}.$$

Thus

$$\mu_2'(n) = 1+\sum_{r=1}^{n-1} rF_r(0)\,F_{n-r-1}(0)+2\sum_{r=1}^{n-1}\sum_{s=1}^{r-1} F_{n-r-1}(0)\,F_{s-1}(0)\,G_{r-s}.$$

Short table of first and second moments about zero, and standard deviation, of the maximum of the partial sums of n independent standard normal deviates

n	μ_1'	μ_2'	σ
3	0·6810	2·1592	1·3021
4	0·9114	2·8842	1·4311
5	1·1108	3·6476	1·5536
6	1·2893	4·4367	1·6657
7	1·4521	5·2446	1·7709
8	1·6029	6·0671	1·8702
9	1·7440	6·9013	1·9647
10	1·8769	7·7451	2·0548
11	2·0031	8·5971	2·1412
12	2·1234	9·4560	2·2242
13	2·2385	10·3210	2·3043
14	2·3492	11·1914	2·3817
15	2·4558	12·0665	2·4567
16	2·5588	12·9459	2·5295
17	2·6585	13·8292	2·6002
18	2·7553	14·7160	2·6691
19	2·8493	15·6060	2·7363
20	2·9409	16·4989	2·8018
21	3·0301	17·3946	2·8659
22	3·1171	18·2928	2·9285
23	3·2022	19·1934	2·9899
24	3·2854	20·0962	3·0500
25	3·3668	21·0010	3·1090

The second term on the right-hand side equals $\frac{1}{2}(n-1)$, by Lemma 3, and the third term reduces, with the aid of Lemma 1, to $\sum_1^{n-2} G_r$. So we finally obtain

$$\mu_2'(n) = \tfrac{1}{2}(n+1)+2\sum_1^{n-2} G_r,$$

where, by (6·11),

$$G_r = \frac{1}{2}\cdot\frac{1}{2\pi}\sum_{j=1}^{r}\{j(r-j+1)\}^{-\frac{1}{2}}.$$

Finally, then $$\mu_2'(n) = \tfrac{1}{2}(n+1) + \frac{1}{2\pi}\sum_{r=1}^{n-2}\sum_{s=1}^{r}\{s(r-s+1)\}^{-\frac{1}{2}}. \qquad (7{\cdot}1)$$

From the point of view of numerical evaluation it is fortunate that the individual terms of the summand in (7·1) are independent of n; computations carried out for a given value of n can be immediately utilized for larger values of n. A short table of specimen values is appended.

Values corresponding to very high n may be approximated by a result of Erdös & Kac (1946), who gave the limiting distribution of $\theta_n = n^{-\frac{1}{2}}U_n$ as

$$\lim_{n\to\infty}\Pr\{\theta_n \leqslant \alpha\} = \sqrt{\frac{2}{\pi}}\int_0^\alpha \exp(-\tfrac{1}{2}x^2)\,dx \quad (\alpha \geqslant 0)$$
$$= 0 \qquad\qquad\qquad\qquad (\alpha < 0).$$

The limiting second moment of θ_n is thus unity, and the asymptotic second moment of U_n is n.

Our results are in agreement with this. We may approximate the double sum in (7·1) by the double integral

$$\int_{y=\frac{1}{2}}^{n-\frac{1}{2}}\int_{x=\frac{1}{2}}^{y} x^{-\frac{1}{2}}(y+1-x)^{-\frac{1}{2}}\,dx\,dy.$$

If we reverse the order of integration, this may easily be evaluated explicitly to give an expression which, to terms of order $n^{\frac{1}{2}}$, reduces to

$$n\pi - 2(2+\sqrt{2})\,n^{\frac{1}{2}}.$$

Thus $$\mu_2' \sim n - \frac{2+\sqrt{2}}{\pi}\sqrt{n}.$$

For $n = 25$ this gives $\mu_2' \sim 19{\cdot}6$; the correct value is $\mu_2' = 21{\cdot}001$.

The author wishes to acknowledge his debt to the referee for the asymptotic results given in the last paragraph.

REFERENCES

ANIS, A. A. & LLOYD, E. H. (1953). On the range of partial sums of a finite number of independent normal variates. *Biometrika*, **40**, 35.

ERDÖS, P. & KAC, M. (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52**, 292.

SPATIAL POINT PROCESSES, WITH APPLICATIONS TO ECOLOGY

By H. R. THOMPSON

Applied Mathematics Laboratory, D.S.I.R., Wellington, New Zealand

1. Introduction

The term 'point processes', referring to stochastic processes in which events occur at more or less irregular intervals and which are represented by points on the time-axis, is of comparatively recent origin, although the existence of such processes has in fact been well known for a long time. They have been discussed fairly extensively in such diverse applications as the counting of radioactive impulses, telephone calls and cases of contagious diseases. Wold (1949) developed a statistical theory for treating processes of this type, and also mentioned briefly how the events could take place in a two-dimensional or higher field. Such a generalization, from events with no time extension to those with no 'space' extension (i.e. specifically of a point character), has a suitable field of application in the ecological study of the distributional pattern of plants. If we can assume to a first approximation that the plants have the dimensions of a point, then we shall see that it is possible to discuss precisely probability relationships between the numbers of plants in different areas of the region under investigation.

The main aims of quantitative ecology are the precise description of a community of plants with interpretations in terms of the biology of the species, and the correlation of vegetational and environmental data, and ecologists have used several methods in an attempt to achieve these aims. In most of the initial work on field sampling for ecological data, the procedure was to take 'quadrats' (sample areas small in relation to the total area of the region) scattered at random over the area, and study statistics derived from the frequency distribution of the numbers of plants per quadrat. While this approach is useful to some extent, in that any given type of distribution function may be fitted to the data, it does not necessarily furnish the kind of information required by an ecologist. It will not give any evidence of trends, or indicate the pattern of the distribution over the area or the way in which this pattern may have arisen, all factors of prime importance in the study of the structure of a plant community. We only have to cite the negative binomial distribution, which is known to arise in at least four different ways, all based on widely differing assumptions, to illustrate this point.

In recent years ecologists have become aware of the need for a more satisfactory approach to the problem, and Greig-Smith (1952) provided a potentially great advance on the statistical side when he recommended the use of a grid of contiguous quadrats over some portion or portions of the region. The advantage, of course, in arranging the quadrats in a grid is that the analysis of variance technique may be employed, either for the detection of trends, or, more importantly, for the detection of a mosaic variation in density (due to ecological causes connected with the spread of the plants) by a 'nested sampling' type of analysis of variance, associating the quadrats into successively larger blocks and comparing the component block variances. The details and applications of this method are described at length by Greig-Smith, together with the results from sampling experiments on artificial

plant communities. We are not concerned here with discussing the ecological implications of the method (for which the reader is referred to a paper to appear elsewhere (Thompson, 1955)), but with the application of point process techniques to deriving the probability relations between the numbers of plants in the quadrats of a grid, required particularly in the analysis of variance and its sampling theory.

Wold's treatment of one-dimensional point processes is based on the distribution of the time interval between successive events, which could be either independent, or dependent on the previous interval, in which case the occurrence of an event would depend on the times of occurrence of the two preceding events. Given $\lambda(x, y)\,dt$, equal to the joint probability that an event has occurred in $(t, t+dt)$ with the two immediately preceding events at instants t_1, t_2, where $x = t-t_1, y = t_1-t_2$, Wold derives integral equations for the distribution functions $F(x, y)$, which is the conditional probability that an event at t_1 is followed by an interval (t_1, t_1+x), given that the immediately preceding event occurred at t_1-y, and $G(x)$, the absolute probability that an event at t_1 is followed by an event in (t_1, t_1+x). From these he obtains the probability $S(\nu, T)$ of exactly ν events in an arbitrary interval of length T. The definition of λ expresses the fact that the process is stationary, for λ remains the same if the set of events undergoes a translation along the time-axis.

This type of approach is adequate for one-dimensional fields because, by means of it, a process in which the course of events depends on all the prehistory up to a given time can be completely specified (apart from random variation); but it does not easily extend to more than one dimension. It seems in fact impossible to specify a stochastic process in two dimensions by means of a similar dependence of events. The difficulty is now (at least in the ecological application) that we are studying a function $F(x, y, T)$ say of the two space variables x, y and the time variable t, for a given value T of t, and we are given, specified by means of their space co-ordinates, a set of events which may have occurred separately at any time. Although the development of the process occurs along a time-axis, with events (i.e. new plants) occurring at specified times, it is mainly the consideration of the spatial pattern with which we are concerned.

For the purpose of finding probability relations between the numbers of plants in neighbouring finite areas, it is convenient to use the method of 'continuous parameter' stochastic processes, developed in connexion with physical problems such as electron cascades, the main contribution being jointly by Bhabha (1950) and Ramakrishnan (1950). They considered the case of particles with specific energies in a continuous range, i.e. effectively point processes in one dimension. Probability relations between the numbers of particles in non-overlapping infinitesimal ranges are defined by density functions of different orders, called by Ramakrishnan 'product densities', and the integrals of these over finite ranges yield linear functions of the product moments of the total numbers of particles in the finite ranges. The required mathematical treatment is obtained by considering a function of the continuous parameter E, $n(E)$ say, which equals 0 everywhere except at a finite number of points $E_1, E_2, \ldots, E_k$ say (at which events occur), where $n(E) = 1$. With suitable modifications this is the approach adopted in this paper, and in the next section the theory of spatial point processes is recapitulated in the most convenient manner. As in Wold, we assume that the processes are stationary in the statistical sense, the essential character being unaltered by any translations of the axes. This is a quite reasonable assumption for plant communities, where trends can often be assumed to be absent.

2. Specification of point stochastic processes in two dimensions

The distribution of plants over a region is described by a continuous parametric system, whose states are labelled by the two continuous variables x and y, the co-ordinates of position of a plant with respect to a given origin. Thus a given pair of values (x, y), which will alternatively be denoted by the parameter A, regarded as a vector, is taken to imply the occurrence of a plant in the infinitesimal rectangle with lower left-hand corner (x, y) and sides of length dx, dy. The infinitesimal element of area $dx\,dy$ will likewise be denoted by the vector dA. A realization of such a system will consist of a number of isolated points, k say, denoted by the parameter values $A_1, A_2, \ldots, A_k$. All states of the system are covered by allowing each A_i to range over the whole domain of A, and letting $k = 0, 1, 2, \ldots$, to give all possible samples. With each possible sample is associated a probability $\Pi(A_1, A_2, \ldots, A_k)$, and a rigorous treatment of continuous parametric systems is obtained (see Bhabha, 1950) by a precise mathematical definition of the probability Π, using set theory. For the purposes of this paper a more direct approach, as given by Ramakrishnan (1950), is sufficient.

Let $N(A)$ denote the number of plants whose positions are below and to the left of the value $A \equiv (x, y)$, i.e. whose parametric values are less than A. Then we can consistently write

$$N(A) = \int^{A} dN(A),$$

and regard $dN(A)$ as denoting the number of plants in the infinitesimal rectangle, area dA, situated at A. Assume that the probability of one plant in dA is $m\,dA$, and the probability of more than one is $o(dA)$. Hence $dN(A)$ is a variable assumed to take only the values 0 or 1, and it follows that all moments of $dN(A)$ are equal to the probability that it takes the value 1. Also the consideration of probability relations between different $dN(A)$ is very much simplified, for a contribution to the product moment $E\{dN(A_1) \ldots dN(A_k)\}$ will only occur when there is a plant in each dA_i, and the contribution is then simply unity. If any one of the dA_i contains no plant, the contribution is zero.

The probability relations between the numbers $dN(A)$ in the small areas dA are defined by functions of the form

$$f_k(A_1, A_2, \ldots, A_k)\,dA_1 dA_2 \ldots dA_k = E\{dN(A_1)\,dN(A_2) \ldots dN(A_k)\}, \qquad (2{\cdot}1)$$

which represents the joint probability that there is one plant in dA_1, one in dA_2, ... one in dA_k, when $dA_1, dA_2, \ldots, dA_k$ are all separate non-overlapping areas. f_k is called a product density of degree k. $f_1(A)\,dA = E\{dN(A)\} = m\,dA$, and it may be noted that the integral of f_1 yields only the mean number of plants in the area of integration, as the addition rule of probabilities does not apply, the events in general not being mutually exclusive. When two of the dA_i are the same, a product density of degree k becomes one of degree $k-1$. For if $dA_k \equiv dA_{k-1}$,

$$\begin{aligned} E\{dN(A_1)\,dN(A_2) \ldots dN(A_{k-1})^2\} &= E\{dN(A_1)\,dN(A_2) \ldots dN(A_{k-1})\} \\ &= f_{k-1}(A_1, A_2, \ldots, A_{k-1})\,dA_1 dA_2 \ldots dA_{k-1}. \end{aligned}$$

It is degeneracies of this type that lead to the following theorem, given by Bhabha, for the product moment of the numbers $N(A_i') - N(A_i)$, denoted by $N_{[i]}$, in finite areas $[A_i, A_i']$ $(i = 1, 2, \ldots, k)$. A_i is the lower left-hand point, A_i' the upper right-hand point of the area,

which is assumed to be rectangular. The k areas denoted by the numbers $[1], [2], \ldots, [k]$, and any number of them may be the same area:

$$E\{N_{[1]}N_{[2]}\ldots N_{[k]}\} = \int_{[1,\ldots,k]} f_1(A)\,dA + \Sigma\int_{[1,\ldots,r]}\int_{[r+1,\ldots,k]} f_2(A_1,A_2)\,dA_1dA_2$$
$$+\Sigma\int_{[1,\ldots,r]}\int_{[r+1,\ldots,s]}\int_{[s+1,\ldots,k]} f_3(A_1,A_2,A_3)\,dA_1dA_2dA_3$$
$$+\ldots+\int_{[1]}\int_{[2]}\ldots\int_{[k]} f_k(A_1,\ldots,A_k)\,dA_1\ldots dA_k, \qquad (2\cdot2)$$

where

(1) by $[1,\ldots,r]$ is meant the equality of the r areas $[1],\ldots,[r]$, if in fact they are equal. If they are not, the terms containing $[1,\ldots,r]$ do not appear in the expression;

(2) the summation sign before each term denotes a summation over like terms in which the integers $1,\ldots,k$ are distributed in all possible ways between the brackets affixed to the integrals, there being no distinction between the order of the integers in a bracket or between the order of the brackets. For example, with $k=4$, integrals of the type of the second term are summed with the groupings of the integers given by

$$[1][234],\quad [2][341],\quad [3][412],\quad [4][123],\quad [12][34],\quad [13][24],\quad [14][23],$$

of the third term by

$$[1][2][34],\quad [1][3][24],\quad [1][4][23],\quad [2][3][14],\quad [2][4][13],\quad [3][4][12].$$

Thus, if the four areas are the same area [1] say, we have for the expectation of the fourth power of the number in [1]

$$E\{N_{[1]}^4\} = \int_{[1]} f_1 + 7\int_{[1]}\int_{[1]} f_2 + 6\int_{[1]}\int_{[1]}\int_{[1]} f_3 + \int_{[1]}\int_{[1]}\int_{[1]}\int_{[1]} f_4,$$

using a condensed notation. If there is more than one area to consider, the expression is modified; for example,

$$E\{N_{[1]}^2N_{[2]}^2\} = \int_{[1]}\int_{[2]} f_2 + \int_{[1]}\int_{[1]}\int_{[2]} f_3 + \int_{[1]}\int_{[2]}\int_{[2]} f_3 + \int_{[1]}\int_{[1]}\int_{[2]}\int_{[2]} f_4.$$

The product densities may consistently be regarded as factorial-moment densities (cf. Bartlett, 1954). For, if we consider the expression for $E\{N_{[1]}^r\}$, the coefficient of the term containing the product density of degree s is the sum of the coefficients of all terms

$$\Sigma\ldots\Sigma\, dN(A_1)^{\alpha_1}dN(A_2)^{\alpha_2}\ldots dN(A_s)^{\alpha_s}$$

for a given s and all different combinations of $\alpha_1,\alpha_2,\ldots,\alpha_s$ subject to the condition $\sum_{i=1}^{s}\alpha_i = r$. That is, the coefficient of $\int_{[1]}\ldots\int_{[1]} f_s$ in $E\{N_{[1]}^r\}$ is $\Sigma r!/(\alpha_1!\,\alpha_2!\ldots\alpha_s!)$ the summation being over all possible different combinations of the $s\alpha_i$. Stevens (1937), in a different connexion, proves that this coefficient is of the general form $\Delta^s0^r/s! = C_s^r$ say, where Δ^s0^r $(s=1,2,\ldots,r)$ is the sth leading difference of the rth powers of the natural numbers. Thus we have

$$E\{N_{[1]}^r\} = C_1^r\int_{[1]} f_1(A)\,dA + C_2^r\int_{[1]}\int_{[1]} f_2(A_1,A_2)\,dA_1dA_2$$
$$+\ldots+C_r^r\int_{[1]}\ldots\int_{[1]} f_r(A_1,\ldots,A_r)\,dA_1\ldots dA_r.$$

One of Stevens's results is, in our notation,

$$N_{[1]}^r = C_1^r N_{[1]} + C_2^r N_{[1]}(N_{[1]}-1) + \ldots + C_r^r N_{[1]}(N_{[1]}-1)\ldots(N_{[1]}-r+1).$$

A simple inductive argument then shows that

$$\int_{[1]}\ldots\int_{[1]} f_k(A_1,\ldots,A_k)\,dA_1\ldots dA_k = E\{N_{[1]}(N_{[1]}-1)\ldots(N_{[1]}-k+1)\}.$$

For a process to be stationary we require that the covariance of the numbers in two areas is independent of their absolute position. Thus, for two areas [1], [2],

$$\begin{aligned}\operatorname{cov}(N_{[1]}, N_{[2]}) &= E\{N_{[1]}-E(N_{[1]})\}\{N_{[2]}-E(N_{[2]})\}\\ &= E\{N_{[1]}N_{[2]}\}-E\{N_{[1]}\}E\{N_{[2]}\}\\ &= \int_{[1]}\int_{[1]}\{f_2(A_1,A_2)-m^2\}\,dA_1\,dA_2, \quad \text{from (2·2).}\end{aligned}$$

This implies that $$f_2(A_1,A_2)-m^2 = \omega(A_2-A_1),$$

so that $f_2(A_1,A_2)$ is a function of the differences x_2-x_1, y_2-y_1.

In practical applications the calculation of the product density of degree k will generally involve a tabulation of the different ways of getting a plant in each of the k infinitesimal areas, due usually to the presence of different classes of plants (different groups or families of plants, different generations, etc.). f_k is then obtained as the sum of the (mutually exclusive) probabilities for all these possible alternative cases, according to the addition rule of probabilities.

3. Analysis of variance on a grid of quadrats

The individual terms of the analysis of variance being considered in this paper are the sums of squares within blocks of a given size and between blocks of the next lowest size. The most useful type of grid to employ is one in which the number of quadrats is a power of 2, leading to blocks of 2, 4, 8, ... quadrats and consequently more information, relative to the size of the grid, than any other arrangement of blocks. The total size of the grid is quite arbitrary; for practical and theoretical purposes it has been fixed here at 256 quadrats, arranged in a 16×16 square. With a suitable choice of quadrat size, the important ecological effects should be found in the terms for the smaller-sized blocks, which have a fairly large number of degrees of freedom and are consequently subject to less fluctuation.

From the definition above, we see that if $B_{k(i)}$ denotes the total number of plants in the ith block of k quadrats, then S_k, the sum of squares within blocks of k quadrats and between blocks of $\frac{1}{2}k$ quadrats, is given by

$$S_k = \frac{2}{k}\sum_{i=1}^{512/k} B_{\frac{1}{2}k(i)}^2 - \frac{1}{k}\sum_{i=1}^{256/k} B_{k(i)}^2, \qquad (3\cdot1)$$

from which S_k has $n_k = 256/k$ degrees of freedom. This is ordinarily the most convenient practical way of calculating individual terms, but for deriving an expected analysis of variance and its sampling errors for any given theoretical model we need to work in terms

of the numbers in a single quadrat. Consider now the moments of S_k for a given k ($k = 2, 4, 8, \ldots, 256$). A three-suffix notation for the quadrat numbers is the most convenient for this purpose.

Let N_{ijl} be the observed number of plants in quadrat (ijl) of the grid, where i ($= 1, 2, \ldots, 256/k$) denotes a primary block of k quadrats, j ($= 1, 2$) a subblock of $\frac{1}{2}k$ quadrats, and l ($= 1, 2, \ldots, \frac{1}{2}k$) a quadrat of a subblock. We assume that blocks of the same size always have identical shapes. For $k = 2^{2p+2}$ ($p = 0, 1, 2, 3$) blocks are taken to be squares of sides 2^{p+1} quadrats, while for $k = 2^{2p+1}$ they are assumed rectangular, of dimensions $2^p \times 2^{p+1}$ quadrats, with the longer length always in the same direction.

From (3·1), transforming $B_{k(i)}$ to a sum of the N_{ijl}, we obtain

$$S_k = \frac{2}{k}\sum_i [(\sum_l N_{i1l})^2 + (\sum_l N_{i2l})^2 - \tfrac{1}{2}(\sum_j \sum_l N_{ijl})^2]$$

$$= \frac{1}{k}\sum_i [\sum_l (N_{i1l} - N_{i2l})]^2. \tag{3·2}$$

From (3·2), by long and tedious but otherwise simple algebra, the higher powers of S_k may be obtained. We quote only S_k^2 in full:

$$\begin{aligned} S_k^2 &= \frac{1}{k^2}[\sum_i \{\sum_l (N_{i1l} - N_{i2l})\}^4 + \sum_{i \neq i'} \{\sum_l (N_{i1l} - N_{i2l})\}^2 \{\sum_l (N_{i'1l} - N_{i'2l})\}^2] \\ &= \frac{1}{k^2}\sum_i \sum_j \{\sum_l N_{ijl}^4 + 4\sum_{l \neq l'} N_{ijl}^3 N_{ijl'} - 4\sum_{l,l'} N_{ijl}^3 N_{ijl'} \\ &\quad + 3\sum_{l \neq l'} N_{ijl}^2 N_{ijl'}^2 + 3\sum_{l,l'} N_{ijl}^2 N_{ijl'}^2 \\ &\quad + 6\sum_{l \neq l' \neq l''} N_{ijl}^2 N_{ijl'} N_{ijl''} + 6\sum_{l' \neq l''} N_{ijl}^2 N_{ijl'} N_{ijl''} \\ &\quad - 12\sum_{l \neq l'} N_{ijl}^3 N_{ijl'} N_{ijl'} + \sum_{l \neq \ldots \neq l''} N_{ijl} N_{ijl'} N_{ijl''} N_{ijl'''} \\ &\quad + 3\sum_{l \neq l',\, l' \neq l''} N_{ijl} N_{ijl'} N_{ijl'} N_{ijl''} - 4\sum_{l \neq l' \neq l''} N_{ijl} N_{ijl'} N_{ijl'} N_{ijl''}\} \\ &\quad + \frac{1}{k^2}\sum_{i \neq i'} \sum_{j,j'} \{\sum_{l,l'} N_{ijl}^2 N_{i'jl'}^2 + 2\sum_{l' \neq l''} N_{ijl}^2 N_{i'jl'} N_{i'jl''} \\ &\quad - 4\sum_{l,l',l''} N_{ijl}^2 N_{i'jl'} N_{i'jl''} + \sum_{l \neq l',\, l' \neq l''} N_{ijl} N_{ijl'} N_{i'jl'} N_{i'jl''} \\ &\quad + 2\sum_{l,l',l''} N_{ijl} N_{ijl'} N_{i'jl'} N_{i'jl''} - 8\sum_{l,l';\, l' \neq l''} N_{ijl} N_{ijl'} N_{i'jl'} N_{i'jl''}\}, \end{aligned} \tag{3·3}$$

where $i, i'; j, j'; l, l', l'', l'''$ are varied independently over their whole respective ranges unless otherwise stated.

For the mean and variance of S_k, we take expected values of (3·2) and (3·3), and see that in the first case

$$E\{S_k\} = \frac{256}{k^2}[kE\{N_{ijl}^2\} + 2\sum_{l \neq l'} E\{N_{ijl} N_{ijl'}\} - 2\sum_{l,l'} E\{N_{ijl} N_{ijl'}\}], \tag{3·4}$$

and is only a linear combination of terms like $E\{N_{[1]}^2\}$, $E[N_{[1]} N_{[2]}\}$, while $E\{S_k^2\}$ will be a linear combination of terms like $E\{N_{[1]}^4\}$, $E\{N_{[1]}^3 N_{[2]}\}$, $E\{N_{[1]}^2 N_{[2]}^2\}$, $E\{N_{[1]}^2 N_{[2]} N_{[3]}\}$, $E\{N_{[1]} N_{[2]} N_{[3]} N_{[4]}\}$.

The expressions for these expectations, obtained from (2·2), are given below, with $f_k(A_1, \ldots, A_k)\, dA_1 \ldots dA_k$ written f_k for short:

$$\left.\begin{aligned}
E\{N_{[1]}^2\} &= \int_{[1]} f_1 + \int_{[1]}\int_{[1]} f_2, \quad E\{N_{[1]}N_{[2]}\} = \int_{[1]}\int_{[2]} f_2,\\
E\{N_{[1]}^4\} &= \int_{[1]} f_1 + 7\int_{[1]}\int_{[1]} f_2 + 6\int_{[1]}\int_{[1]}\int_{[1]} f_3 + \int_{[1]}\int_{[1]}\int_{[1]}\int_{[1]} f_4,\\
E\{N_{[1]}^3 N_{[2]}\} &= \int_{[1]}\int_{[2]} f_2 + 3\int_{[1]}\int_{[1]}\int_{[2]} f_3 + \int_{[1]}\int_{[1]}\int_{[1]}\int_{[2]} f_4,\\
E\{N_{[1]}^2 N_{[2]}^2\} &= \int_{[1]}\int_{[2]} f_2 + \int_{[1]}\int_{[1]}\int_{[2]} f_3 + \int_{[1]}\int_{[2]}\int_{[2]} f_3 + \int_{[1]}\int_{[1]}\int_{[2]}\int_{[2]} f_4,\\
E\{N_{[1]}^2 N_{[2]} N_{[3]}\} &= \int_{[1]}\int_{[2]}\int_{[3]} f_3 + \int_{[1]}\int_{[1]}\int_{[2]}\int_{[3]} f_4,\\
E\{N_{[1]} N_{[2]} N_{[3]} N_{[4]}\} &= \int_{[1]}\int_{[2]}\int_{[3]}\int_{[4]} f_4.
\end{aligned}\right\} \tag{3·5}$$

4. Application of the theory to an ecological model

The simplest assumption mathematically that can be made about a community of plants is that they are distributed at random in the Poisson distribution. Such a distribution might conceivably arise when an area is first invaded by wind-borne seeds, but it is hardly ever encountered in practice. A more realistic and more useful model is obtained by allowing plants distributed in this manner, i.e. at random, to become the centres of distribution ('parents') of a generation of 'offspring' plants, whose positions depend on those of their respective parents. The offspring of a given parent are assumed to be distributed independently of each other, the distances r from parent to offspring following the isotropic normal distribution

$$f(r)\, dr = e^{-\frac{1}{2}r^2/\sigma^2}\, dr/(2\pi\sigma^2), \tag{4·1}$$

and the numbers n of offspring per parent following an arbitrary probability law $p(n)$. We consider for simplicity the distribution of the offspring only, for then the only distances which have to be taken into account are those between offspring of the same group. and these are also isotropic normal, but with parameter $\sqrt{2}\,\sigma$. For, $f(r)\, dr$ is the product of two independent components $e^{-\frac{1}{2}x^2/\sigma^2}\, dx/\sqrt{(2\pi\sigma^2)}$, $e^{-\frac{1}{2}y^2/\sigma^2}\, dy/\sqrt{(2\pi\sigma^2)}$ along rectangular axes, and the distance between two offspring projected on either of these axes is merely the sum of two independent distances (from the parent) following a one-dimensional normal distribution.

The product density f_k of degree k is found (§ 2) as the sum of mutually exclusive probabilities, denoted by $\Pr(dA_1, \ldots, dA_k)$, for all the possible alternatives with k different areas $dA_1, \ldots, dA_k$. Here the contributions to f_k will arise from all different combinations of k plants from s groups ($s \leqslant k$) such that there are α_1 plants from the first group, ..., α_s from the sth with $\sum_{i=1}^{s} \alpha_i = k$. Of these we need only consider the case $s = 1$, i.e. all k plants from the same group; the other cases will merely result in products of lower order product densities since the contributions from different groups are entirely independent. We

prove now that the joint probability of k offspring of the same group occurring in $dA_1, \ldots, dA_k$ is

$$\Pr(dA_1, \ldots, dA_k) = m_0 E\{n(n-1)\ldots(n-k+1)\}\,\phi(A_1, \ldots, A_k)\,dA_1 \ldots dA_k, \tag{4·2}$$

where

$$\left.\begin{aligned}\phi(A_1, \ldots, A_k) &= \left(\frac{1}{2\pi\sigma^2}\right)^{k-1}\frac{1}{k}\exp\left[-\sum_{i=1}^{k-1}\sum_{j>i}\frac{R_{ij}^2}{2k\sigma^2}\right],\\ R_{ij}^2 &= (x_i-x_j)^2+(y_i-y_j)^2,\end{aligned}\right\} \tag{4·3}$$

and m_0 is the mean density of the parent plants. The mean density of offspring plants is $m = m_0 E\{n\}$.

Assume that the position of the parent plant of a group with a random number n of offspring is at dA_0. The probability of this is

$$m_0 dA_0.$$

The probability of k ($\leqslant n$) offspring of this parent in $dA_1, \ldots, dA_k$, distant $R_{01}, \ldots, R_{0k}$ from dA_0, given the parents' position is at dA_0, is

$$n(n-1)\ldots(n-k+1)\left(\frac{1}{2\pi\sigma^2}\right)^k \exp\left[-\sum_{i=1}^{k}\frac{R_{0i}^2}{2\sigma^2}\right] dA_1 \ldots dA_k,$$

since there are n possibilities for dA_1, $n-1$ for $dA_2, \ldots, n-k+1$ for dA_k, and each distance is distributed independently. Therefore, averaging over n,

$$\begin{aligned}&\Pr(dA_0, dA_1, \ldots, dA_k)\\ &\quad = \Pr(dA_0)\Pr(dA_1, \ldots, dA_k \mid dA_0)\\ &\quad = m_0 E\{n(n-1)\ldots(n-k+1)\}\left(\frac{1}{2\pi\sigma^2}\right)^k \exp\left[-\sum_{i=1}^{k}\frac{(x_0-x_i)^2+(y_0-y_i)^2}{2\sigma^2}\right] dA_0 dA_1 \ldots dA_k\\ &\quad = m_0 E\{n(n-1)\ldots(n-k+1)\}\frac{1}{2\pi\sigma^2}\exp\left[-k\frac{(x_0-\bar{x})^2+(y_0-\bar{y})^2}{2\sigma^2}\right] dA_0\\ &\qquad\qquad \times\left(\frac{1}{2\pi\sigma^2}\right)^{k-1}\exp\left[-\sum_{i=1}^{k-1}\sum_{j>i}^{k-1}\frac{(x_i-x_j)^2+(y_i-y_j)^2}{2k\sigma^2}\right] dA_1 \ldots dA_k,\end{aligned}$$

where

$$\bar{x} = \sum_{i=1}^{k}\frac{x_i}{k}, \quad \bar{y} = \sum_{i=1}^{k}\frac{y_i}{k}.$$

Integrating out with respect to $dA_0 (= dx_0 dy_0)$ over the whole region (assumed infinite), we obtain (4·2). $\phi(A_1, \ldots, A_k)$ may be shown to be a product of two $(k-1)$-variate normal distributions (in x and y), the $k-1$ variables being, for example, the differences of the first $k-1$ values of x (or y) from the kth, so that each variable is distributed normally, mean zero, variance $2\sigma^2$, and the correlation between any two variables is $\frac{1}{2}$.

For the mean values and variances of the set of sums of squares we require the product densities of second, third and fourth degree. With two areas dA_1 and dA_2, there are only two possibilities to consider; both areas might contain plants from different groups, in which case the joint probability $\Pr(dA_1, dA_2)$ is $m^2 dA_1 dA_2$, or else the plants may be offspring from the same group. We note that the case where one of the dA contains no plant is

automatically excluded, since it makes no contribution to the second degree product density. For the second case, we have from (4·2)

$$\Pr(dA_1, dA_2) = m_0 E\{n(n-1)\}\,\phi(A_1, A_2)\,dA_1 dA_2,$$

so that, adding the two mutually exclusive probabilities, we obtain

$$\begin{aligned} f_2(A_1, A_2)\,dA_1 dA_2 &= [m_0 E\{n(n-1)\}\,\phi(A_1, A_2) + m^2]\,dA_1 dA_2 \\ &= m\left(\frac{g_2}{4\pi\sigma^2} e^{-R_{12}^2/4\sigma^2} + m\right) dA_1 dA_2, \end{aligned} \tag{4·4}$$

from (4·3) where

$$g_r = E\{n(n-1)\ldots(n-r+1)\}/E\{n\}. \tag{4·5}$$

Now tabulate the possibilities for three plants in three different areas dA_1, dA_2, dA_3. If the plants are all in different groups,

$$\Pr(dA_1, dA_2, dA_3) = m^3 dA_1 dA_2 dA_3.$$

If dA_1, dA_2 are in a group G, dA_3 not in G, then since the contributions from different groups are independent, we have

$$\Pr(dA_1, dA_2, dA_3) = m_0 E\{n(n-1)\}\,\phi(A_1, A_2)\,dA_1 dA_2 . m\,dA_3.$$

If dA_1, dA_2, dA_3 are all in the same group,

$$\Pr(dA_1, dA_2, dA_3) = m_0 E\{n(n-1)(n-2)\}\,\phi(A_1, A_2, A_3)\,dA_1 dA_2 dA_3.$$

This exhausts all the possibilities; therefore,

$$\begin{aligned} f_3(A_1, A_2, A_3) &= m^3 + m^2 g_2\{\phi(A_1, A_2) + \phi(A_1, A_3) + \phi(A_2, A_3)\} + m g_3 \phi(A_1, A_2, A_3) \\ &= m^3 + m^2 g_2 (e^{-R_{12}^2/4\sigma^2} + e^{-R_{13}^2/4\sigma^2} + e^{-R_{23}^2/4\sigma^2})/4\pi\sigma^2 \\ &\quad + m g_3\, e^{-(R_{12}^2 + R_{13}^2 + R_{23}^2)/6\sigma^2}/12\pi^2\sigma^4. \end{aligned} \tag{4·6}$$

We omit the tabulation for the fourth degree case, as it is obvious from the expression

$$\begin{aligned} f_4(A_1, A_2, A_3, A_4) = m^4 &+ m^3 g_2\{\phi(A_1, A_2) + \phi(A_1, A_3) + \phi(A_1, A_4) + \phi(A_2, A_3) \\ &+ \phi(A_2, A_4) + \phi(A_3, A_4)\} \\ &+ m^2 g_2^2\{\phi(A_1, A_2)\,\phi(A_3, A_4) + \phi(A_1, A_3)\,\phi(A_2, A_4) \\ &+ \phi(A_1, A_4)\,\phi(A_2, A_3)\} \\ &+ m^2 g_3\{\phi(A_1, A_2, A_3) + \phi(A_1, A_2, A_4) + \phi(A_1, A_3, A_4) + \phi(A_2, A_3, A_4)\} \\ &+ m g_4 \phi(A_1, A_2, A_3, A_4). \end{aligned} \tag{4·7}$$

The product densities are integrated and substituted in (3·5) to give the expectations needed in the calculation of $E\{S_k\}$, $E\{S_k^2\}$. An explicit solution has been found for the integral of f_2 only, but product densities of all degrees may be integrated numerically. Let the quadrats of the grid be squares of side h, and write

$$\int_{[1]}\int_{[2]}\ldots\int_{[k]} \phi(A_1, \ldots, A_k)\,dA_1 \ldots dA_k = h^2 I_{12\ldots k}, \tag{4·8}$$

where square brackets round each suffix of I may be assumed if desired. Then equations (3·5) become

$$
\left.\begin{aligned}
E\{N_{[1]}^2\} &= mh^2 + mh^2 g_2 I_{11}, \quad E\{N_{[1]}N_{[2]}\} = mh^2 g_2 I_{12},\\
E\{N_{[1]}^4\} &= mh^2 + 7m^2h^4 + 6m^3h^6 + m^4h^8 + mh^2(7 + 18mh^2 + 6m^2h^4)\, g_2 I_{11}\\
&\quad + 3m^2h^4 g_2^2 I_{11}^2 + mh^2(6 + 4mh^2)\, g_3 I_{111} + mh^2 g_4 I_{1111},\\
E\{N_{[1]}^3 N_{[2]}\} &= m^2h^4 + 3m^3h^6 + m^4h^8 + 3m^2h^4(1 + mh^2)\, g_2 I_{11} + m^2h^4 g_3 I_{111}\\
&\quad + mh^2(1 + 6mh^2 + 3m^2h^4)\, g_2 I_{12} + 3m^2h^4 g_2^2 I_{11} I_{12}\\
&\quad + 3mh^2(1 + mh^2)\, g_3 I_{112} + mh^2 g_4 I_{1112},\\
E\{N_{[1]}^2 N_{[2]}^2\} &= m^2h^4 + 2m^3h^6 + m^4h^8 + 2m^2h^4(1 + mh^2)\, g_2 I_{11}\\
&\quad + mh^2(1 + 4mh^2 + 4m^2h^4)\, g_2 I_{12} + m^2h^4 g_2^2 (I_{11}^2 + 2I_{12}^2)\\
&\quad + mh^2(1 + 2mh^2)\, g_3 (I_{112} + I_{122}) + mh^2 g_4 I_{1122},\\
E\{N_{[1]}^2 N_{[2]} N_{[3]}\} &= m^3h^6 + m^4h^8 + m^3h^6 g_2 I_{11} + (m^2h^4 + 2m^3h^6)\, g_2 (I_{12} + I_{13})\\
&\quad + (m^2h^4 + m^3h^6)\, g_2 I_{23} + m^2h^4 g_2^2 (I_{11} I_{23} + 2I_{12} I_{13})\\
&\quad + (mh^2 + 2m^2h^4)\, g_3 I_{123} + m^2h^4 g_3 (I_{112} + I_{113}) + mh^2 g_4 I_{1123},\\
E\{N_{[1]} N_{[2]} N_{[3]} N_{[4]}\} &= m^4h^8 + m^3h^6 g_2 (I_{12} + I_{13} + I_{14} + I_{23} + I_{24} + I_{34})\\
&\quad + m^2h^4 g_2^2 (I_{12} I_{34} + I_{13} I_{24} + I_{14} I_{23})\\
&\quad + m^2h^4 g_3 (I_{123} + I_{124} + I_{134} + I_{234}) + mh^2 g_4 I_{1234}.
\end{aligned}\right\} \quad (4\cdot9)
$$

The details of the calculations are omitted, as much tedious manipulation is involved.

Table 1. *Expected mean squares for the non-random model*

k	2	4	8	16	32	64	128	256
n_k	128	64	32	16	8	4	2	1
$E\{S'_k/mh^2\}$	1·12	1·19	1·64	1·84	2·40	2·58	2·91	3·02

There are many different numerical examples possible, since there are two sets of parameters to be varied, those of $p(n)$ and those of $f(r)$. A study of the effects of different functions $p(n)$ and $f(r)$ on the form of the distribution of $N_{[i]}$, the number of plants in a single quadrat, has already been made (Thompson, 1954). Table 1 gives the expected mean squares $E\{S'_k/mh^2\}$, $(S'_k = S_k/n_k)$, for the particular case $\sigma = h/\sqrt{2}$ (which gives a simplification of the numerical integrations) and a binomial distribution for $p(n)$ with a mean of three offspring per group. The parameters in $p(n) = \binom{N}{n} p^n (1-p)^{N-n}$ $(n = 0, 1, \ldots, N)$ are $N = 6$, $p = \frac{1}{2}$, and $g_r = (N-1)(N-2)\ldots(N-r+1)\, p^{r-1}$. The set of expected mean squares $E\{S'_k\}$ has upper and lower limits, reached as σ tends to 0 and ∞ respectively; they are

$mh^2(1+g_2)$ and mh^2, independent of k. These are both mean squares for a Poisson distribution, and are to be expected, since in the first case each offspring's position coincides with that of its parent, and in the second case there is effectively no clumping.

The standard error of a sum of squares is not in itself a very useful statistic, because of the skew distribution of S_k, at least for a small number of degrees of freedom. For the first three mean squares of Table 1. with fairly large n_k (and for which normal approximations might be expected to hold), the standard errors are, respectively, 0·20, 0·27, 0·46 for $mh^2 = 1$, and 0·13, 0·16, 0·26 for $mh^2 = \frac{1}{2}$. A more useful practical method of studying the sampling variation is to determine the limits of error of individual mean squares, so that the set of expected mean squares for a given model has appropriate to it 'significance bands' outside which terms observed from samples of the model would only be expected to fall with a given (small) probability. In this approach we ignore the fact that a real correlation exists between any pair of mean squares, except for the case when the $N_{[i]}$ are independent normal variables, and calculate limits of error on the assumption of approximate χ^2 distributions for individual mean squares, based on their first two moments, since we know that for $N_{[i]}$ normal (variance σ^2 say) S_k/σ^2 is distributed exactly as χ^2 with n_k degrees of freedom. Table 2 gives these limits of error for the case $mh^2 = 1$ only, but to a first approximation they may be taken to apply to mean squares $S'_k/(mh^2)$ so long as mh^2 is not too different from unity. The bands are symmetrical, in the sense that the probability of an observed mean square being above the upper limit equals the probability of its falling below the lower limit.

Table 2. *Approximate p % significance bands for the mean squares of Table* 1 ($mh^2 = 1$)

k	2	4	8	16	32	64	128	256
95 % band	0·76 1·55	0·72 1·78	0·86 2·65	0·72 3·50	0·62 5·38	0·28 7·28	0·06 10·83	0·00 15·22
80 % band	0·86 1·39	0·86 1·55	1·07 2·25	0·99 2·78	0·96 4·08	0·62 5·08	0·26 6·72	0·03 8·18

It is instructive to compare these figures (applying to a model with not a very high degree of contagion) with the results for a purely random distribution (parents alone, say), which may be shown to approximate very closely to the normal case. The correlations between individual mean squares are in this random case almost negligible, and S_k/mh^2 follows almost exactly a χ^2 distribution with $n_k\{2kmh^2/(1+2kmh^2)\}$ degrees of freedom. For a Poisson distribution, we have explicitly

$$\left.\begin{aligned} E\{S'_k\} &= mh^2, \\ \operatorname{var}\{S'_k\} &= mh^2(1+2kmh^2)/256. \end{aligned}\right\} \tag{4·10}$$

Table 3 gives the limits of error for the case $mh^2 = 1$, but as in Table 2 they may equally well apply for other values of mh^2. In fact, for $mh = \frac{1}{2}, \infty$, which are symmetrically placed with respect to $mh^2 = 1$ here, the differences from the values given in the tables are never more than $\pm 0{\cdot}02$.

Table 3. *Approximate p % significance bands for the mean squares from a Poisson distribution* ($mh^2 = 1$)

k	2	4	8	16	32	64	128	256
95 % band	0·75 1·29	0·67 1·40	0·56 1·56	0·42 1·82	0·27 2·20	0·12 2·79	0·03 3·70	0·00 5·03
80 % band	0·83 1·18	0·77 1·25	0·69 1·34	0·58 1·48	0·43 1·68	0·26 1·95	0·10 2·30	0·02 2·71

5. DISCUSSION

When an analysis of variance is carried out on a set of observational data, the usual aim is to test the homogeneity of the set of variances obtained, employing the F test. With a grid of quadrats, the appropriate method would be to test the ratio $(S_k/n_k)/(S_2/n_2)$, which, on the null hypothesis that the numbers $N_{[i]}$ in the quadrats are independently and normally distributed with the same variance, follows the F distribution with n_2 and n_k degrees of freedom. The conditions for the applicability of this test will not, however, generally hold in the present ecological application, for the associations of plants in clumps result in the non-independence of the numbers in neighbouring quadrats, and even more distant quadrats if the clumps themselves are related, and the effect of clumping is usually to produce a contagious distribution of the quadrat numbers. Under the usual null hypothesis, H_0 say, we have

$$\Pr\{F \geqslant F_\alpha \mid H_0\} = \alpha, \tag{5·1}$$

where $F = (S_k/n_k)/(S_2/n_2)$, and F_α is the value of F on n_2 and n_k degrees of freedom at the 100α % level of significance. The expected value of F on this hypothesis is unity; however, for most non-random models, including the example above, it is greater. (It may occasionally be less than unity.) Therefore, to study the power function of the F test we should ideally calculate the probability

$$\Pr\{F/E(F) \geqslant F_\alpha \mid H\} = \beta, \tag{5·2}$$

where H is the non-normal hypothesis relating to the particular model, and compare β with α.

This has been done approximately for the example of § 4, for the two cases $k = 4, 8$ and with $\alpha = 0{\cdot}05, 0{\cdot}01$, the results being given in Table 4. The method used is one recommended by David & Johnson (1951). Rewrite (5·2) in the form

$$\Pr\{S_k - aS_2 \geqslant 0 \mid H\} = \beta, \tag{5·3}$$

where $a = n_k F_\alpha E(F)/n_2 = 2F_\alpha E(F)/k$. The first two moments of $S = S_k - aS_2$ are sufficient to find this probability approximately, transforming S so that probability levels for the χ^2 distribution can be used. A new statistic $S/(mh^2) + C$ is formed whose first two moments are exactly those of a χ^2 distribution with f degrees of freedom by taking

$$C = \tfrac{1}{2}\operatorname{var}\{S/(mh^2)\} - E\{S/(mh^2)\}, \quad f = \tfrac{1}{2}\operatorname{var}\{S/(mh^2)\}.$$

Then $\Pr\{\chi^2_{(f)} \geqslant C\}$ furnishes β. The method is highly satisfactory when the $N_{[i]}$ are Poisson or normal variables. These results show that the actual level of significance being used is

Table 4. *Values of* β *(from* (5·3)*) for the model of* § 4

	$\alpha = 0{\cdot}05$			$\alpha = 0{\cdot}01$		
k \ mh^2	$\frac{1}{2}$	1	∞	$\frac{1}{2}$	1	∞
4	0·090	0·073	0·055	0·033	0·023	0·013
8	0·117	0·103	0·089	0·048	0·038	0·028

quite close to its assumed true value. However, we have assumed on *a priori* knowledge of $E\{F\}$, which would not normally be so in practice. A truer picture of what would happen if the F test were applied indiscriminately is given in Table 5, where we have calculated

$$\Pr\{F \geqslant F_\alpha \mid H\} = \beta'. \tag{5·4}$$

In this particular example, we see that the use of the F test on unadjusted variances is of no value at all for k greater than 4, and even for $k = 4$ the differences between β' and α are becoming appreciable.

Table 5. *Table of* β' *(from* (5·4)*), for the model of* § 4

	$\alpha = 0{\cdot}05$			$\alpha = 0{\cdot}01$		
k \ mh^2	$\frac{1}{2}$	1	∞	$\frac{1}{2}$	1	∞
4	0·139	0·119	0·098	0·054	0·040	0·027
8	0·535	0·537	0·539	0·350	0·342	0·334

We have only discussed one particular example in this paper. Several others giving rise to non-random distributions of plants have been derived, consistent with the mathematical and computational difficulty involved. They are described at length in Thompson (1955), with more emphasis on the ecological application, however, so that it may be of interest to note some of them briefly here. We can extend the model of § 4 to include several generations of plants, the distances from each offspring to its parent still following the isotropic normal distribution (4·1). The offspring of the original parent become the parents of another generation of offspring, which in their turn become parent plants, and so on. The distance between any two plants of the same group also follows (4·1), but with parameter depending on the number of direct steps between the two plants. If the original parents are randomly distributed, we find similarly to (4·4),

$$f_2(A_1, A_2) - m^2 = m_0 \sum_\alpha E\{2n_\alpha\} e^{-R_{12}^2/2\alpha\sigma^2}/2\pi\alpha\sigma^2, \tag{5·5}$$

where n_α is the number of pairs of plants in the group which have the distances between them distributed with parameter $\sqrt{\alpha}\,\sigma$.

Models using the isotropic normal distribution for $f(r)$ produce over-dispersion of individuals (ratio of variance to mean in a single quadrat greater than unity). If the x and y components of the distance between successive plants in a chain follow independent χ^2 distributions with $2f$ degrees of freedom ($f \geqslant 2$), then under-dispersion of individuals results because of the 'negative correlation' effect introduced (cf. Bartlett, 1954). The frequency function of the distance between plants g generations apart is easily obtained, being simply the product of two independent χ^2 distributions with $2fg$ degrees of freedom. For a group with n plants (the nth being the offspring of the $(n-1)$th, and so on), the original parent being randomly distributed and with $f = 2$, we have

$$f_2(A_1, A_2) - m^2 = m_0 \sum_{g=1}^{n-1} 2(n-g)(\lambda_1\lambda_2)^{2g}[(x_2-x_1)(y_2-y_1)]^{2g-1} e^{-\lambda_1(x_2-x_1)-\lambda_2(y_2-y_1)}/[(2g-1)!]^2, \tag{5·6}$$

where λ_1, λ_2 are arbitrary and are considered in relation to the quadrat size h, and $x_2 \geqslant x_1 \geqslant 0$, $y_2 \geqslant y_1 \geqslant 0$. This represents approximately an actual field example, in which a definite forward move is made in each generation due to vegetative spreading of the plants, and the probability of having two successive plants very close together is small. The other models discussed are developments of these two basic types and have no intrinsic mathematical interest. In all of them, however, an attempt has been made to keep the ecological aspect in mind so that they should be descriptive in some way of an idealized plant community.

I wish to acknowledge gratefully my indebtedness to Prof. M. S. Bartlett for suggesting this subject of research, and for his advice and assistance during my investigations at Manchester University. I also wish to thank the New Zealand Department of Scientific and Industrial Research for financial assistance at that time.

REFERENCES

BARTLETT, M. S. (1954). Processus stochastiques ponctuels. *Ann. Inst. H. Poincaré*, **14**, 35.

BHABHA, H. J. (1950). On the stochastic theory of continuous parametric systems and its application to electron cascades. *Proc. Roy. Soc.* A, **202**, 301.

DAVID, F. & JOHNSON, N. L. (1951). The effect of non-normality on power function of the F test in the analysis of variance. *Biometrika*, **38**, 43.

GREIG-SMITH, P. (1952). The use of random and contiguous quadrats in the study of the structure of plant communities. *Ann. Bot.*, N.S., **16**, 293.

RAMAKRISHNAN, A. (1950). Stochastic processes relating to particles distributed in a continuous infinity of states. *Proc. Camb. Phil. Soc.* **46**, 595.

STEVENS, W. L. (1937). Significance of grouping. *Ann. Eugen., Lond.*, **8**, 57.

THOMPSON, H. R. (1954). A note on contagious distributions. *Biometrika*, **41**, 268.

THOMPSON, H. R. (1955). Statistical study of plant distribution patterns using a grid of contiguous quadrats (in the Press).

WOLD, H. (1949). Sur les processus stationnaires ponctuels. Le calcul des probabilités et des applications. *Publications du C.N.R.S.*, **13**.

THE OUTCOME OF A STOCHASTIC EPIDEMIC—A NOTE ON BAILEY'S PAPER

By P. WHITTLE

Applied Mathematics Laboratory, New Zealand Department of Scientific and Industrial Research

In a recent paper (1953) N. Bailey has considered a stochastic epidemic model of the type set up by Bartlett (1949), and has shown that the probability distribution (P_w) of the ultimate number of infected individuals (w) may be calculated by solving a certain set of doubly recurrent relations. I propose to show that for quite a general case these same probabilities may be obtained by the solution of a set of singly recurrent relations (eqs. (17) and (24)). Furthermore, an expression may be derived (eqs. (38) and (40)) for the probability that an infection introduced into a large population will 'take'—this provides a stochastic equivalent to Kermack & McKendrick's threshold theorem (1927 and later).

1. Introductory

Following Bailey, we shall assume that an initial a infectious cases are introduced into a population of n uninfected but susceptible individuals, and that the probability that after a time t there are r susceptibles still uninfected and s infectious cases not yet removed is $p_{rs}(t)$. Let $Bs\Delta t$ be the probability that an infectious individual is removed in the infinitesimal time interval $(t, t+\Delta t)$, and let $A_r s\Delta t$ be the corresponding probability that a new infection takes place. No particular form is assumed for the function A_r at the moment.

The development of the probabilities p_{rs} is then governed by the relations

$$\frac{\partial}{\partial t} p_{rs} = A_{r+1}(s-1)\, p_{r+1,s-1} + B(s+1)\, p_{r,s+1} - (A_r s + Bs)\, p_{rs}. \tag{1}$$

which become

$$(\lambda + A_r s + Bs)\, q_{rs} = A_{r+1}(s-1)\, q_{r+1,s-1} + B(s+1)\, q_{r,s+1} + \delta_{ar}\delta_{as} \tag{2}$$

if the transformation

$$q_{rs} = \int_0^\infty e^{-\lambda t} p_{rs}(t)\, dt \tag{3}$$

is performed. As Bailey observes, the probability of an epidemic of total size w is then

$$P_w = B f_{n-w,1}, \tag{4}$$

where

$$f_{rs} = \lim_{\lambda \to 0} q_{rs}. \tag{5}$$

This limit exists as long as $s > 0$, and the equations regulating the appropriate f_{rs} are obtained simply by setting λ equal to zero in equation (2), with $s = 1, 2, \ldots$. These are, in effect, the equations used by Bailey for the computation of the P_w.

2. Establishment of the recurrence relations

If we write

$$h_{r0} = 0, \quad h_{rs} = s f_{rs} \quad (s = 1, 2, \ldots), \tag{6}$$

$$\alpha_r = \frac{B}{A_r + B}, \tag{7}$$

$$\beta_r = \frac{A_{r+1}}{A_r + B}, \tag{8}$$

then the equations for the f_{rs} take the form

$$h_{rs} = \alpha_r h_{r,s+1} + \beta_r h_{r+1,s-1}, \tag{9}$$

$$h_{ns} = \alpha_n h_{n,s+1} + \beta_n \left(\frac{\delta_{as}}{A_{n+1}}\right) \quad (r = n-1, n-2, \ldots;\ s = 1, 2, \ldots). \tag{10}$$

Now, let

$$H_r(x) = \sum_{s=1}^{\infty} h_{rs} x^{s+1}. \tag{11}$$

Multiplying (9) by x^{s+1} and summing over s we find

$$H_r(x) = \frac{x^2}{x - \alpha_r} [\beta_r H_{r+1}(x) - \alpha_r h_{r1}]. \tag{12}$$

A direct solution of (9) shows that

$$h_{r1} = \frac{\beta_r}{\alpha_r} H_{r+1}(\alpha_r), \tag{13}$$

as, indeed, it must, if the expression (12) for H_r is to constitute a finite series in x. We have thus

$$H_r(x) = \frac{\beta_r x^2}{x - \alpha_r} [H_{r+1}(x) - H_{r+1}(\alpha_r)], \tag{14}$$

a relation which certainly holds for $r = n-1, n-2, \ldots$, and also for $r = n$ if we introduce a function

$$H_{n+1}(x) = \frac{x^a}{A_{n+1}} \tag{15}$$

(cf. eq. (10)). Further, by (4), (6) and (13) we have

$$P_w = B \frac{\beta_{n-w}}{\alpha_{n-w}} H_{n-w+1}(\alpha_{n-w}). \tag{16}$$

The required probabilities P_w can thus be derived by solving the simply recurrent relation (14) with initial condition (15), and then using (16). By doing almost precisely this we shall obtain an explicit recurrence relation for the P_w. With the help of (14) and (15) $H_{n-w+1}(x)$ can be expressed in terms of $H_{n-w+2}(\alpha_{n-w+1}) \ldots H_n(\alpha_{n-1})$, $H_{n+1}(x)$. Setting $x = \alpha_{n-w}$ in this expression and substituting for the $H_{r+1}(\alpha_r)$ from (16) we find

$$\sum_{w=0}^{u} \kappa_{n-u+1,n-u} \kappa_{n-u+2,n-u} \ldots \kappa_{n-w,n-u} \frac{\alpha_{n-w}}{\beta_{n-w}} P_w = \kappa_{n-u+1,n-u} \kappa_{n-u+2,n-u} \ldots \kappa_{n,n-u} \left(\frac{B}{A_{n+1}}\right) \alpha_{n-u}^a, \tag{17}$$

where

$$\kappa_{rr} = 1, \quad \kappa_{rs} = \frac{\alpha_s^2 \beta_r}{\alpha_s - \alpha_r} \quad (r \neq s). \tag{18}$$

The final relation of (17), for $u = n$, reduces to

$$\sum_{0}^{n} P_w = 1, \tag{19}$$

provided that A_0 is zero, as it must be.

Complications arise if any of the A_r's (and consequently the corresponding α_r's) are equal. When this happens P_w involves not only $H_{n+1}(x)$ but also its derivatives. The extreme case in this direction is that for which A_r is constant for $r > 0$

$$\left.\begin{aligned} A_r &= A \quad (r = 1, 2, \ldots), \\ A_0 &= 0, \end{aligned}\right\} \tag{20}$$

as is approximately the case during the initial stages of an epidemic if the number of susceptibles is large. The appropriate modification of (17) may be derived by a repeated application of de l'Hôpital's rule, and may be shown by induction to have a solution

$$\left.\begin{aligned} P_w &= \frac{A^w B^{a+w}}{(A+B)^{a+2w}} \frac{a(a+2w-1)!}{w!\,(a+w)!} \\ P_n &= 1-\sum_0^{n-1} P_w \end{aligned}\right\} \quad (w = 0, 1, \ldots, n-1). \tag{21}$$

(Formula (21) is a generalization of one obtained by D. G. Kendall in a slightly different context, see formula (52) of his paper of 1948.) We shall return to this case later, and shall for the moment consider the more usual alternative

$$A_r = Cr, \tag{22}$$

where C is a constant. We have

$$\left.\begin{aligned} \alpha_r &= \frac{B}{B+Cr}, \qquad \beta_r = \frac{C(r+1)}{B+Cr}, \\ \kappa_{rs} &= \frac{B(r+1)}{(B+Cs)\,(r-s)} = \frac{\alpha_s(r+1)}{r-s}, \end{aligned}\right\} \tag{23}$$

and (17) becomes

$$\sum_{w=0}^{u} \binom{n-w}{n-u} \alpha_{n-u}^{-w} P_w = \binom{n}{u} \alpha_{n-u}^{a} \quad (u = 0, 1, \ldots, n). \tag{24}$$

For computational purposes it is convenient to consider instead of P_w the quantity

$$Q_w = \frac{(n-w)!}{n!} P_w, \tag{25}$$

for which

$$\sum_{w=0}^{n} \frac{\alpha_{n-u}^{-w} Q_w}{(u-w)!} = \frac{\alpha_{n-u}^{a}}{u!}. \tag{26}$$

3. The probability of epidemic

Let us now return to expression (21). We shall use this as a comparison formula for establishing the behaviour of more refined models. Note, however, that the model upon which it is based is a perfectly valid one, which for quite large ranges of w is more realistic than that corresponding to assumption (22), since this assumption requires that the population mix homogeneously, a requirement never fulfilled in a large population.

The following table gives the first few values of P_w as calculated from formulae (21) and (24) for the case $a = 2$, $n = 30$, $\rho = B/C = 30B/A = 10$:

	w	0	1	2	3
P_w	$A_r = 3$	0·0625	0·0234	0·0110	0·0058
	$A_r = 0{\cdot}1r$	0·0625	0·0251	0·0122	0·0078

The difference between the two sets of probabilities is small but increases with increasing w.

We shall now adopt the following definition: It shall be said that an epidemic has (has not) taken place if the total proportion of susceptibles which become infected exceeds (does not exceed) a predetermined fraction γ. With this definition the probability of no epidemic is

$$\pi_\gamma = \sum_{w=0}^{n\gamma} P_w, \tag{27}$$

where P_w is in general given by (17).

Consider now two other models for which the infection intensities are given by A'_r, A''_r. We shall assume that in all three cases the infection intensity is non-decreasing with increasing r:

$$\left.\begin{aligned} A_{r+1} &\geqslant A_r \\ A'_{r+1} &\geqslant A'_r \\ A''_{r+1} &\geqslant A''_r \end{aligned}\right\} \quad (r = 0, 1, \ldots). \tag{28}$$

We shall further assume that, at least for the range $n \geqslant r \geqslant n(1-\gamma)$, the intensity for the first model lies uniformly between those for the other two

$$A'_r \geqslant A_r \geqslant A''_r. \tag{29}$$

It is then intuitively evident that

$$\sum_0^{n\gamma} P'_w \leqslant \sum_0^{n\gamma} P_w \leqslant \sum_0^{n\gamma} P''_w. \tag{30}$$

Suppose now that the intensities for those two comparison models have the constant values

$$\left.\begin{aligned} A'_r &= A_n \\ A''_r &= A_{n(1-\gamma)} \end{aligned}\right\} \quad (r > 0), \tag{31}$$

while $A'_0 = A''_0 = 0$. Condition (29) is thus fulfilled, and the inequality (30) becomes

$$\sum_0^{n\gamma} S_w(A_n) \leqslant \pi_\gamma \leqslant \sum_0^{n\gamma} S_w(A_{n(1-\gamma)}), \tag{32}$$

where we have used $S_w(A)$ to denote the expression in (21).

Consider now the partial sum $\sum_0^{n\gamma} S_w(A)$ as n becomes large. We have

$$\frac{S_{w+1}}{S_w} = \frac{AB}{(A+B)^2} \frac{(a+2w+1)(a+2w)}{(w+1)(a+w+1)} < \frac{4AB}{(A+B)^2} \quad \text{if} \quad w > \frac{(a+1)(a-4)}{6}$$

$$= 4k \quad \text{(say)}. \tag{33}$$

The quantity $4k$ will be less than unity, except for the case $A = B$, which we shall exclude for the moment. We can thus write

$$\sum_0^{n\gamma} S_w(A) = \sum_0^{\infty} S_w(A) - R_{n\gamma}(A), \tag{34}$$

where

$$R_{n\gamma}(A) = \sum_{n\gamma+1}^{\infty} S_w(A) < S_{n\gamma+1}(A)\,[1 + (4k) + (4k)^2 + \ldots] = O[(4k)^{n\gamma}]. \tag{35}$$

The infinite sum in (34) has the value

$$\sum_0^{\infty} S_w(A) = \left[\frac{1-\sqrt{(1-4k)}}{2k}\right]^a \left(\frac{B}{A+B}\right)^a = \left[\frac{A+B-|A-B|}{2A}\right]^a. \tag{36}$$

Combining (32), (34), (35) and (36) we have thus

$$\left[\frac{A_n+B-|A_n-B|}{2A_n}\right]^a+O[(4k)^{n\gamma}]\leqslant\pi_\gamma\leqslant\left[\frac{A_{n(1-\gamma)}+B-|A_{n(1-\gamma)}-B|}{2A_{n(1-\gamma)}}\right]^a. \tag{37}$$

For large n the remainder term in the first member of inequality (37) is quite negligible, and we shall no longer include it. Note now from (36) that we have $\Sigma S_w(A)=1$ or $(B/A)^a$ according as A is less or greater than B. Setting these evaluations in (37) we find that there are at least three distinct cases:

$$\left.\begin{aligned} &A_n>B,\ A_{n(1-\gamma)}>B: && (B/A_n)^a\leqslant\pi_\gamma\leqslant(B/A_{n(1-\gamma)})^a,\\ &A_n>B,\ A_{n(1-\gamma)}<B: && (B/A_n)^a\leqslant\pi_\gamma\leqslant 1,\\ &A_n<B,\ A_{n(1-\gamma)}<B: && \pi_\gamma=1. \end{aligned}\right\} \tag{38}$$

We may sum up the situation in terms of Bailey's removal ratio

$$\rho_n=B/C=nB/A_n, \tag{39}$$

the ratio of removal and infection rates for a population of size n:

$$\left.\begin{aligned} &\text{For } \rho_n<n \text{ and } \rho_{n(1-\gamma)}<n(1-\gamma), \text{ the probability of epidemic lies between}\\ &1-\left(\frac{\rho_n}{n}\right)^a \quad\text{and}\quad 1-\left(\frac{\rho_{n(1-\gamma)}}{n(1-\gamma)}\right)^a,\\ &\text{For } \rho_n<n \text{ and } \rho_{n(1-\gamma)}>n(1-\gamma), \text{ the probability of epidemic lies between}\\ &\text{zero} \quad\text{and}\quad 1-\left(\frac{\rho_n}{n}\right)^a,\\ &\text{For } \rho_n>n \text{ and } \rho_{n(1-\gamma)}>n(1-\gamma), \text{ the probability of epidemic is zero.} \end{aligned}\right\} \tag{40}$$

These statements provide an equivalent in the stochastic case to Kermack & McKendrick's threshold theorem in the deterministic case, at least for the case of large populations. Since for large n the ratio $\rho_n/n=B/A_n$ will tend to an almost constant value, statements (40) may be roughly condensed to

$$\left.\begin{aligned} &\text{For } \rho_n<n \text{ the probability of epidemic is } 1-(\rho_n/n)^a,\\ &\text{For } \rho_n>n \text{ the probability of epidemic is zero.} \end{aligned}\right\} \tag{41}$$

The transition case $\rho_n=n$ cannot be adequately treated by approximate considerations of the present type. Equation (28) indicates that the probability of completion is of fairly constant magnitude for small w, roughly of order $(\frac{1}{2})^a$. As the number of susceptibles diminish, however, the critical value of ρ will fall, and it seems likely that the epidemic will eventually be halted, although only after having made appreciable inroads on the population.

The statements (36) are reminiscent of similar statements concerning the natural extinction of populations (cf. Bartlett, 1946) and could have been derived by regarding the group of infected persons as a population with birth and death rates A/n and B. To reason in this way is unsatisfactory, however, since the condition that the infected group shall ultimately disappear provides no guarantee that infection will be confined to a preassigned fraction of the population of susceptibles.

The argument above is probably more illuminating in the following intuitive form. The probability distribution P_w of (24) presents two different forms according as A is less than

or greater than B (see Fig. 1). In both cases P_w dwindles with increasing w, and if the population size is large enough to permit w to take large values P_w will finally approach zero. In the case $A < B$ the sum of the P_w up to this stage will approach unity. In the case $A > B$ this sum will have some value less than unity, say $1-\alpha$, so that P_n must have a finite value α if relation (19) is to be fulfilled.

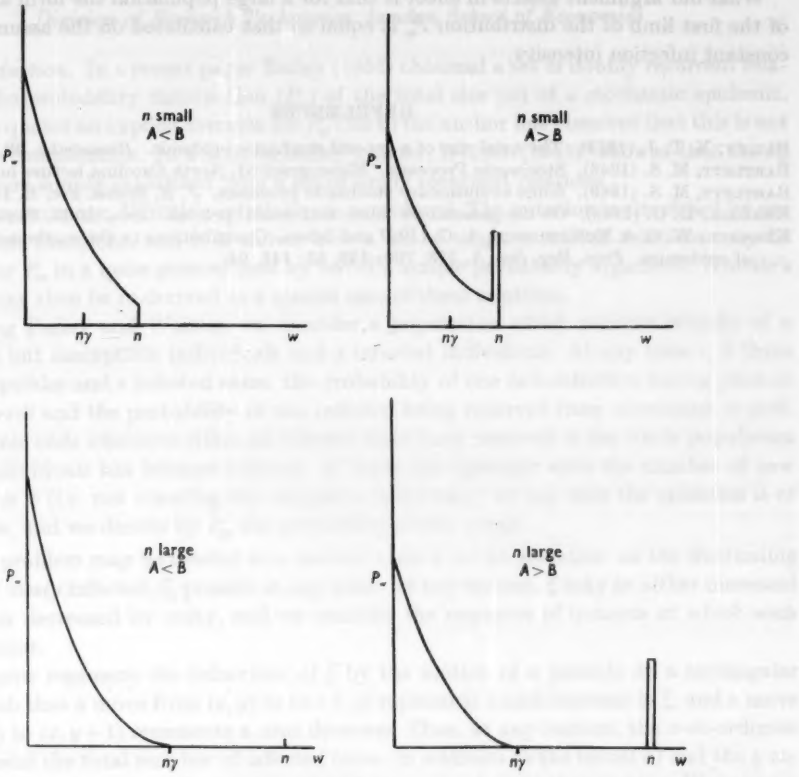


Fig. 1. The appearance of the distribution curve P_w of the ultimate number of infected persons, on the assumption of a constant infection rate.

For large n the probability of no epidemic

$$\pi_\gamma = \sum_0^{n\gamma} S_w(A) \tag{42}$$

will be equal to the area under the initial part of the curve: unity if $A < B$, α if $A > B$.

The fact that all probability mass which does not fall in the first J-shaped part of the curve falls at $w = n$ indicates that either the epidemic keeps within bounds (probability $1-\alpha$) or else it infects the entire population (probability α).

Models with varying A show a similar, although less extreme behaviour. Thus the distribution curves calculated by Bailey are either J-shaped or U-shaped, depending upon the relative values of the removal ratio and the population size. The J-shaped curves

correspond to cases in which the infection is almost certainly confined to a small proportion of the population. The U-shaped curves correspond to cases in which the infection strikes at *either* a small proportion *or* a large proportion of the population, the probabilities of these two alternatives being equal to the integrals of the corresponding limbs of the probability distribution.

What our argument asserts in effect is that for a large population the form and integral of the first limb of the distribution P_w is equal to that calculated on the assumption of a constant infection intensity.

REFERENCES

BAILEY, N. T. J. (1953). The total size of a general stochastic epidemic. *Biometrika*, **40**, 177.
BARTLETT, M. S. (1946). Stochastic Processes. Mimeographed, North Carolina lecture notes.
BARTLETT, M. S. (1949). Some evolutionary stochastic processes. *J. R. Statist. Soc.* B, **11**, 211.
KENDALL, D. G. (1948). On the generalized 'birth-and-death' process. *Ann. Math. Statist.* **19**, 1.
KERMACK, W. O. & MCKENDRICK, A. G. (1927 and later). Contributions to the mathematical theory of epidemics. *Proc. Roy. Soc.* A, **115**, 700; **138**, 55; **141**, 94.

A NOTE ON BAILEY'S AND WHITTLE'S TREATMENT OF A GENERAL STOCHASTIC EPIDEMIC

By F. G. FOSTER

Division of Research Techniques, London School of Economics

1. *Introduction.* In a recent paper Bailey (1953) obtained a set of doubly recurrent relations for the probability distribution (P_w) of the total size (w) of a stochastic epidemic. Bailey also quotes an explicit formula for P_w due to the author but observes that this is not suitable for computation. In a note on Bailey's paper, Whittle (1955)* showed that the P_w could be more simply calculated from a set of singly recurrent relations.

In the present note, Whittle's relations are considered. The notation is simplified by use of symmetric functions, and it is shown how a set of singly recurrent relations may be obtained for P_w in a quite general case by use of a simple probability argument. Whittle's relations may then be re-derived as a special case of these relations.

Following Bailey and Whittle, we consider a population which consists initially of n uninfected but susceptible individuals and a infected individuals. At any time t, if there are r susceptibles and s infected cases, the probability of one new infection taking place in time dt is $rs\,dt$ and the probability of one infected being removed from circulation is $\rho s\,dt$. The epidemic ends whenever either all infected have been removed or the whole population of $a+n$ individuals has become infected. If when the epidemic ends the number of new infections is w (i.e. not counting the original a infections), we say that the epidemic is of total size w, and we denote by P_w the probability of this event.

2. The problem may be treated as a random walk if we fix attention on the fluctuating number of *newly* infected, ξ, present at any time. At any instant, ξ may be either increased by unity or decreased by unity, and we consider the sequence of instants at which such changes occur.

Let us now represent the behaviour of ξ by the motion of a particle on a rectangular lattice, such that a move from (x, y) to $(x+1, y)$ represents a unit increase in ξ, and a move from (x, y) to $(x, y+1)$ represents a unit decrease. Thus, at any instant, the x-co-ordinate will represent the total number of infected cases (in addition to the initial a) and the y-co-ordinate the total number of removals which has occurred up to that instant. We suppose that the probabilities of the moves from (x, y) to $(x+1, y)$ and to $(x, y+1)$ are respectively λ_x, μ_x ($\lambda_x+\mu_x=1$), which in general depend on x. For example, in Bailey's problem:

$$\lambda_x = \frac{n-x}{n-x+\rho}, \quad \mu_x = \frac{\rho}{n-x+\rho}.$$

The motion starts at $(0, 0)$, and stops as soon as any point of the barriers,

$$x = n, \quad y = x+a,$$

is reached. This corresponds to the end of the epidemic, and we are interested in the probability P_w that the particle stops on $(w, a+w)$ ($w = 0, 1, \ldots, n-1$). In the general case we define

$$P_n \equiv 1 - \sum_{w=0}^{n-1} P_w.$$

* See pp. 116–22 above of the present issue.

Thus we have a Birth-and-Death process with the parameters λ_x, μ_x, and the novelty resides in the fact that these parameters are position-dependent. It will be noted, however, that they are assumed independent of the y-co-ordinate. We study the problem for quite general parameters of this type.

3. We now introduce some notation for symmetric functions. Denote by $h_l(x,m)$ the homogeneous product-sum of weight l in the $(m+1)$ quantities, μ_x, μ_{x+1}, ..., μ_{x+m} (cf. MacMahon, 1915). When these quantities are assumed distinct, define

$$h_l^{(j)}(x,m) = \mu_{x+j}^{l+m}/(\mu_{x+j}-\mu_x)\dots(\mu_{x+j}-\mu_{x+j-1})(\mu_{x+j}-\mu_{x+j+1})\dots(\mu_{x+j}-\mu_{x+m}).$$

Then we have the formula $$h_l(x,m) = \sum_{j=0}^{m} h_l^{(j)}(x,m).$$

Now define $$q_l(x,m) = \lambda_x\lambda_{x+1}\dots\lambda_{x+m-1}h_l(x,m),$$

and $$q_l^{(j)}(x,m) = \lambda_x\lambda_{x+1}\dots\lambda_{x+m-1}h_l^{(j)}(x,m).$$

In this notation, it may be verified that Whittle's set of recurrence relations becomes

$$r_k: \quad \sum_{i=0}^{k} q_i^{(i)}(k-i,i)\,P_{k-i} = q_{a+k}^{(k)}(0,k) \quad (k=0,1,\dots,n-1).$$

4. As Whittle observes, in these relations we have to assume that the μ's are all distinct. We proceed now to derive the general formula, using a simple probability argument.

When no barriers are present, the probability that the particle attains the point $(x+m, y+l)$ from the point (x,y) is $q_l(x,m)$. Now P_k is the probability of attaining $(k, a+k)$ from $(0,0)$ *by any path below the barrier*. Therefore

$$P_k = q_{a+k}(0,k) - q_k(0,k)\,P_0 - q_{k-1}(1,k-1)\,P_1 \dots - q_1(k-1,1)\,P_{k-1}.$$

We may rewrite this as the set of relations,

$$\rho_k: \quad \sum_{i=0}^{k} q_i(k-i,i)\,P_{k-i} = q_{a+k}(0,k) \quad (k=0,1,\dots,n-1).$$

This formula is valid quite generally, and Whittle's formula may be obtained from it by use of the following relations, which are easily verified:

$$\rho_k = \sum_{i=0}^{k} q_i^{(0)}(k-i,i)\,r_{k-i} \quad (k=0,1,\dots,n-1).$$

5. As an example, we consider the special case where the λ's and μ's are constant:

$$\lambda_x \equiv \lambda, \quad \mu_x \equiv \mu.$$

Then $$q_r(s,t) = \lambda^t\mu^r\binom{r+t}{t}.$$

Thus $$\sum_{i=0}^{k} (\lambda\mu)^i\binom{2i}{i} P_{k-i} = \lambda^k\mu^{a+k}\binom{a+2k}{k} \quad (k=0,1,\dots,n-1),$$

which has the solution $$P_k = \lambda^k\mu^{a+k}\frac{a}{a+2k}\binom{a+2k}{k}.$$

It is interesting to note the connexion between the stochastic epidemic problem and results related to the arc sine law in the theory of fluctuations in coin-tossing (cf. Feller, 1950, p. 252). Thus the case of constant λ, μ may be applied to the tossing of a biased coin with probabilities λ, μ of heads or tails. Then P_k is the probability that the number of heads is greater than or equal to the number of tails plus a for the first time at the $(2k+a)$th trial. The probability

$$P_n = 1 - \sum_{k=0}^{n-1} P_k$$

is interpreted as the probability that in $(2n+a-1)$ trials the number of heads is always less than the number of tails plus a. For the particular case $a = 1$, $\lambda = \mu = \frac{1}{2}$, we have thus

$$P_n = 1 - \sum_{k=0}^{n-1} \frac{1}{2^{2k+1}} \frac{1}{2k+1} \binom{2k+1}{k}.$$

and it may be verified that this equals $2^{-2n}\binom{2n}{n}$, in agreement with the formula of Theorem 1 in Feller (1950, p. 252).

REFERENCES

BAILEY, NORMAN T. J. (1953). The total size of a stochastic epidemic. *Biometrika*, **40**, 177.

FELLER, WILLIAM (1950). *An Introduction to Probability Theory and its Applications*. New York: John Wiley and Sons.

MACMAHON, P. A. (1915). *Combinatory Analysis*. Cambridge University Press.

WHITTLE, P. (1955). The outcome of a stochastic epidemic—a note on Bailey's paper. *Biometrika*, **42**, 116–22.

THE DETERMINISTIC MODEL OF A SIMPLE EPIDEMIC FOR MORE THAN ONE COMMUNITY

By S. RUSHTON and A. J. MAUTNER
Imperial College, London

1. Introduction

In general, a deterministic model can be expected to give a satisfactory picture of a developing process such as the spread of an epidemic, or the growth of a population, so long as the numbers of individuals are sufficiently large. On the other hand, if the numbers are small the continuous description is no longer valid; moreover, the effects of chance occurrences become appreciable in any particular instance. so that a probability treatment is necessary. The literature on deterministic and stochastic models for epidemics has been reviewed by Bailey (1952): to the references given there should be added Bailey 1953*a*, *b*). Much of this previous work refers essentially to a single community or to a set of isolated communities. In the present note we consider the deterministic model of a simple epidemic for several related communities. A simple epidemic is one in which it is assumed that infection spreads only by contact between individuals, and that none of the infected individuals is removed from circulation by death, recovery or isolation. We do not suggest that the model we discuss here is a realistic one in the sense that any actual epidemic has followed precisely this over-simplified pattern, but all epidemic models so far discussed have lacked realism in some sense. We do suggest, however, that the model discussed in this note can provide a basis for further work along these particular lines so that more realistic and therefore more complicated models can be treated by starting from our basic solution.

2. The differential equations for m communities

We denote by $y_i(t)$ the number of susceptibles in the ith community at time t and by n_i the total size of the ith community ($i = 1, \ldots, m$). Then, on the assumptions (i) that there is homogeneous mixing within each community with α_i as the internal infection rate in the ith community and (ii) that there is also homogeneous mixing between communities with β_{ij} as the infection rate between the ith and jth communities, the differential equations for the y_i's are

$$\frac{dy_i}{dt} = -y_i\{\alpha_i(n_i - y_i) + \sum_{j \neq i} \beta_{ij}(n_j - y_j)\} \quad (i = 1, \ldots, m). \tag{1}$$

Under the simplifying assumptions that $n_i = n$ and $\alpha_i = \alpha$ (all i), and that $\beta_{ij} = \alpha\gamma$ (all $i \neq j$) and taking αt as the time-scale and continuing to denote it by t, these equations become

$$\frac{dy_i}{dt} = -y_i\{(n - y_i) + \gamma \sum_{j \neq i} (n - y_j)\} \quad (i = 1, \ldots, m). \tag{2}$$

The parameter γ in these equations is then the ratio of the assumed common 'cross-infection rate' between communities to the common internal infection rate α, which takes account only of internally generated infection. The solutions of the set (2) are completely

determined by m independent initial conditions, e.g. the m values $y_i(0)$ $(i = 1, \ldots, m)$, and it follows that if any set of these initial values are equal then the corresponding functions $y_i(t)$ are identical for all t. In particular, if all the $y_i(0)$ are equal, the set (2) reduce to a single equation

$$dy/dt = -(1+(m-1)\gamma)(n-y)y$$

for $y(t) = y_1(t) = \ldots = y_m(t)$. This is just the equation for a simple epidemic in a single community with total infection rate $\alpha(1+(m-1)\gamma)$.

Another particular case, whose solution is considered below in §§ 4 and 5, occurs when $(m-1)$ of the initial values $y_i(0)$ are equal and the remaining one, say $y_1(0)$, is not equal to the others. For example, this case arises if a single infected person appears initially in only one of the communities (say the first), so that

$$y_1(0) = n-1 \quad \text{and} \quad y_2(0) = y_3(0) = \ldots = y_m(0) = n.$$

The equations (2) in such a case reduce to the pair of equations

$$\left.\begin{aligned} dy_1/dt &= -y_1(n-y_1+(m-1)\gamma(n-y_2)), \\ dy_2/dt &= -y_2(\gamma(n-y_1)+(1+(m-2)\gamma)(n-y_2)), \end{aligned}\right\} \tag{3}$$

since $y_2(t) = y_3(t) = \ldots = y_m(t)$ for all t.

3. A GENERAL SOLUTION

If the initial values $y_i(0)$ $(i = 1, \ldots, m)$ are all distinct we can obtain a solution of the equations (2) as follows. Writing equations (2) in the form

$$\frac{dy_i}{dt} = y_i(y_i+\gamma \sum_{j \neq i} y_j - na) \quad (i = 1, \ldots, m), \tag{4}$$

where $a = 1+(m-1)\gamma$, and making the transformations

$$y_i = nY_i, \quad s = nt, \quad Y_i = e^{-as}U_i, \tag{5}$$

we obtain the equations

$$\frac{1}{U_i}\frac{dU_i}{ds} = e^{-as}(U_i+\gamma \sum_{j \neq i} U_j) \quad (i = 1, \ldots, m). \tag{6}$$

Changing the independent variable to

$$v = (1-e^{-as})/a, \tag{7}$$

these equations become

$$\frac{1}{U_i}\frac{dU_i}{dv} = \frac{d}{dv}\log U_i = U_i+\gamma \sum_{j \neq i} U_j \quad (i = 1, \ldots, m). \tag{8}$$

Denoting the matrix of coefficients on the right of these equations by $\mathbf{A}$, the vector of elements U_i by $\mathbf{U}$, and denoting a column vector by $\{\ \}$, this set is

$$\frac{d}{dv}\{\log U_i\} = \mathbf{AU}. \tag{9}$$

It follows that

$$\frac{d}{dv}\mathbf{A}^{-1}\{\log U_i\} = \frac{d}{dv}\left\{\log \prod_{j=1}^{m} U_j^{a^{ij}}\right\} = \mathbf{U}, \tag{10}$$

where, corresponding to $\mathbf{A} = [a_{ij}]$, we denote $\mathbf{A}^{-1} = [a^{ij}]$.

Now, putting
$$X_i = \prod_{j=1}^{m} U_j^{a_{ij}}, \tag{11}$$
so that
$$U_i = \prod_{j=1}^{m} X_j^{a_{ij}}, \tag{12}$$
the equations (10) become
$$\frac{d}{dv}\log X_i = \frac{1}{X_i}\frac{dX_i}{dv} = \prod_{j=1}^{m} X_j^{a_{ij}} = \left(\prod_{j=1}^{m} X_j\right)^{\gamma} X_i^{1-\gamma} \tag{13}$$
for $i = 1, \ldots, m$. We then have
$$X_i^{\gamma-2}\frac{dX_i}{dv} = \left(\prod_{j=1}^{m} X_j\right)^{\gamma} = F(v), \quad \text{say,} \tag{14}$$
for all i. Integrating these equations we obtain
$$X_i = (\alpha_i - G)^{-1/(1-\gamma)} \quad (i = 1, \ldots, m), \tag{15}$$
where
$$G(v) = (1-\gamma)\int_0^v F(v)\,dv, \tag{16}$$
and
$$\alpha_i = (X_i(0))^{\gamma-1}. \tag{17}$$
From (12) and the definition of $F(v)$ in (14) we have
$$U_i = \left(\prod_{j=1}^{m} X_j\right)^{\gamma} X_i^{1-\gamma} \tag{18}$$
$$= F(v)/(\alpha_i - G) \quad (i = 1, \ldots, m) \tag{19}$$
from (15).

Since, from (14), (15) and (16),
$$F(v) = \left\{\prod_{i=1}^{m} (\alpha_i - G)\right\}^{-\gamma/(1-\gamma)} = \frac{1}{1-\gamma}\frac{dG}{dv}, \tag{20}$$
we may obtain an integral expression for v as
$$v = \frac{1}{1-\gamma}\int_0^G \left\{\prod_{i=1}^{m} (\alpha_i - G)\right\}^{\gamma/(1-\gamma)} dG. \tag{21}$$
By means of (19) and (21), using the expression for $F(v)$ in terms of G in (20), the solutions for U_i and the time parameter v are given parametrically in terms of G.

4. The special case of equations (3)

In the special case mentioned at the end of § 2 in which $y_2(0) = \ldots = y_m(0)$ and $y_1(0) \neq y_2(0)$, the general equations (2) reduce to the pair of equations (3).

We see from (5) and (18) that these conditions imply that
$$U_2(0) = U_3(0) = \ldots = U_m(0), \quad U_1(0) \neq U_2(0)$$
and
$$X_2(0) = X_3(0) = \ldots = X_m(0), \quad X_1(0) \neq X_2(0).$$
Defining
$$(X_1(0))^{\gamma-1} = \alpha, \tag{22}$$
$$(X_i(0))^{\gamma-1} = \beta \quad \text{for} \quad i = 2, 3, \ldots, m, \tag{23}$$
it follows from (16) that
$$U_1 = F(v)/(\alpha - G) \tag{24}$$
and
$$U_i = F(v)/(\beta - G) \quad (i = 2, \ldots, m). \tag{25}$$

Now put $$Z = U_1/U_2, \quad W = U_2 - U_1, \tag{26}$$

so that $$U_2 = W/(1-Z), \quad U_1 = WZ/(1-Z), \tag{27}$$

and note that if $y_2(0) > y_1(0)$, which we may take to be the case, then $y_2(t) > y_1(t)$ for all t, and it follows that $U_2 > U_1$, so that $W \geqslant 0$ and $0 \leqslant Z \leqslant 1$ for all t. Also we must have $\alpha > \beta$.

Expressing W in terms of Z by means of the above equations and the expression for $F(v)$ in (20), we obtain $$W = A\{(1-Z)^{\mu}/Z^{\mu-1}\}^{\lambda}, \tag{28}$$

where $\lambda = 1/(1-\gamma)$, $\mu = 2+(m-2)\gamma$ and A is an arbitrary constant $= (\alpha-\beta)^{\lambda(1-\mu)}$.

From (20) we have $$\frac{dv}{dG} = \frac{\lambda}{F(v)}.$$

It is easily seen that this is equivalent to

$$\frac{dv}{dZ} = -\frac{\lambda}{ZW}, \tag{29}$$

so that, corresponding to (21), we may express v in terms of Z as

$$v = B - \frac{\lambda}{A}\int_0^Z \{(1-Z)^{\mu}/Z^{\mu-1}\}^{-\lambda}\, dZ/Z, \tag{30}$$

where B is a second arbitrary constant. The integral in (30) is an Incomplete Beta Function which can be expressed in terms of the hypergeometric function. In practice, however, it is preferable to evaluate the integral in (30) directly.

5. Practical form of the solution

Considering the practical form of the solution of § 4, we note that if we choose a new time-parameter $$T = (1+(m-1)\gamma)\, v = (1 - e^{-(1+(m-1)\gamma)s}), \tag{31}$$

where we recall that we have put $s = nt$, then $0 \leqslant T \leqslant 1$ and the whole history of the epidemic is shown on a standard time scale extending over a unit interval. Equations (28) and (30) constitute the solution to our problem, the solution being expressed parametrically in terms of $Z = U_1/U_2 = Y_1/Y_2 = y_1/y_2$.

Defining $$g(Z) = \{(1-Z)^{\mu}/Z^{\mu-1}\}, \tag{32}$$

then, from (28), (30) and (31),

$$W = Ag(Z), \tag{33}$$

and $$T = C - \lambda(\mu-1)\int_0^Z (1/WZ)\, dZ, \tag{34}$$

$$= C - (\lambda(\mu-1)/A)\int_0^Z dZ/(Zg(Z)). \tag{35}$$

The initial values of W and Z are

$$W_0 = U_2(0) - U_1(0) = Y_2(0) - Y_1(0) = (y_2(0) - y_1(0))/n, \tag{36}$$

and $Z_0 = y_1(0)/y_2(0)$. We notice incidentally that if $y_2(0) = n$, so that initially there are no infected persons in any community other than the first, then $Z_0 = 1 - W_0$.

The arbitrary constants A and C are clearly given by

$$A = W_0/g(Z_0) \tag{37}$$

and $$C = (\lambda(\mu-1)/A)\int_0^{Z_0} dZ/(Zg(Z)). \tag{38}$$

From equations (5) and (31) the proportions Y_1 and Y_2 of susceptibles in the communities at any time are

$$\begin{aligned} Y_2 &= (1-T)\,U_2 \\ &= (1-T)\,W/(1-Z) = A(1-T)\,g(Z)/(1-Z) \end{aligned} \tag{39}$$

in virtue of (27), and

$$Y_1 = ZY_2. \tag{40}$$

Table 1. *Numerical procedure*: $\gamma = 0{\cdot}2$, $m = 2$, $Z_0 = 0{\cdot}95$

Z	$g(Z)$	$\int_0^Z \frac{dZ}{Zg(Z)}$	T	s	Y_2	Y_1	$-\frac{dY_2}{ds}$	$-\frac{dY_1}{ds}$	$-\left(\frac{dY_1}{ds}+\frac{dY_2}{ds}\right)$
0·95	$5{\cdot}96\times10^{-4}$	57·03	0·000	0·00	1·000	0·950	0·010	0·047	0·057
0·94	$9{\cdot}53\times10^{-4}$	42·94	0·252	0·24	0·996	0·937	0·016	0·060	0·077
0·93	$1{\cdot}42\times10^{-3}$	33·73	0·417	0·45	0·992	0·923	0·023	0·073	0·096
0·92	$2{\cdot}01\times10^{-3}$	27·32	0·531	0·63	0·987	0·908	0·031	0·086	0·117
0·91	$2{\cdot}73\times10^{-3}$	22·65	0·615	0·80	0·981	0·893	0·039	0·099	0·138
0·90	$3{\cdot}61\times10^{-3}$	19·13	0·678	0·94	0·975	0·877	0·049	0·112	0·161
0·89	$4{\cdot}64\times10^{-3}$	16·40	0·727	1·08	0·968	0·862	0·057	0·123	0·181
0·88	$5{\cdot}85\times10^{-3}$	14·23	0·765	1·21	0·959	0·844	0·069	0·138	0·207
0·87	$7{\cdot}25\times10^{-3}$	12·47	0·797	1·33	0·951	0·827	0·080	0·151	0·231
0·86	$8{\cdot}85\times10^{-3}$	11·03	0·823	1·44	0·941	0·809	0·092	0·164	0·256
0·85	$1{\cdot}07\times10^{-2}$	9·83	0·844	1·55	0·930	0·791	0·104	0·177	0·280
0·84	$1{\cdot}27\times10^{-2}$	8·81	0·862	1·66	0·915	0·769	0·120	0·191	0·310
0·83	$1{\cdot}50\times10^{-2}$	7·95	0·878	1·75	0·907	0·753	0·129	0·200	0·330
0·82	$1{\cdot}76\times10^{-2}$	7·20	0·891	1·85	0·893	0·733	0·143	0·212	0·355
0·81	$2{\cdot}05\times10^{-2}$	6·56	0·903	1·94	0·880	0·713	0·157	0·222	0·379
0·80	$2{\cdot}36\times10^{-2}$	5·99	0·913	2·03	0·865	0·692	0·170	0·232	0·402
0·75	$4{\cdot}48\times10^{-2}$	3·99	0·949	2·47	0·774	0·580	0·240	0·270	0·510
0·70	$7{\cdot}70\times10^{-2}$	2·81	0·970	2·91	0·654	0·458	0·297	0·280	0·577
0·65	$1{\cdot}24\times10^{-1}$	2·05	0·983	3·41	0·500	0·325	0·318	0·252	0·569
0·60	$1{\cdot}92\times10^{-1}$	1·53	0·992	4·08	0·301	0·181	0·260	0·173	0·433
0·55	$2{\cdot}87\times10^{-1}$	1·16	0·999	5·89	0·046	0·025	0·052	0·029	0·081

In terms of Y_1 and Y_2 and the time parameter s $(0 \leqslant s \leqslant \infty)$, where

$$s = \{\log_e(1-T)\}/(1+(m-1)\gamma), \tag{41}$$

the epidemic rate $-dY_1/ds$ in the first community and $-dY_2/ds$ in the other communities are given by

$$\left.\begin{aligned} dY_1/ds &= -Y_1(Y_1+(m-1)\gamma Y_2-(1+(m-1)\gamma)), \\ dY_2/ds &= -Y_2(\gamma Y_1+(1+(m-2)\gamma)\,Y_2-(1+(m-1)\gamma)), \end{aligned}\right\} \tag{42}$$

and the total epidemic rate is $-(dY_1/ds+(m-1)\,dY_2/ds)$.

For communities of total size n the actual epidemic rates are just $-dy_i/dt = -n^2\,dY_i/ds$.

The numerical procedure in applying the solution is illustrated in Table 1 for the simple case of $m = 2$ communities with $\gamma = 0{\cdot}2$ and initially 5 % of infected individuals in the first community and none in the second. In this case $Y_1(0) = 0{\cdot}95, Y_2(0) = 1, W_0 = 0{\cdot}05, Z_0 = 0{\cdot}95$,

$\lambda = \frac{5}{4}$ and $\mu = 2$. The arbitrary constants are $A = 83{\cdot}8712$ and $C = 1{\cdot}0199$. The principal part of the numerical work is the computing of $g(Z)$ and the integral $\int_0^Z dZ/(Zg(Z))$. In this example, the epidemic is effectively completed by the time that Z has decreased in value to 0·55, when $T = 0{\cdot}999$, and the values of $g(Z)$ and $\int_0^Z dZ/(Zg(Z))$ are therefore only shown for $0{\cdot}95 \geqslant Z \geqslant 0{\cdot}55$. If the epidemic rates $-dY_1/ds$ and $-dY_2/ds$ are plotted the curve for $-dY_2/ds$ is at first below that for $-dY_1/ds$ as we should expect, crosses the latter at about its maximum, rises to a slightly higher maximum and remains above the curve for $-dY_1/ds$ until the completion of the epidemic. The curve for the total epidemic rate falls somewhat more slowly from its maximum than the rate at which it rises. We have examined a number

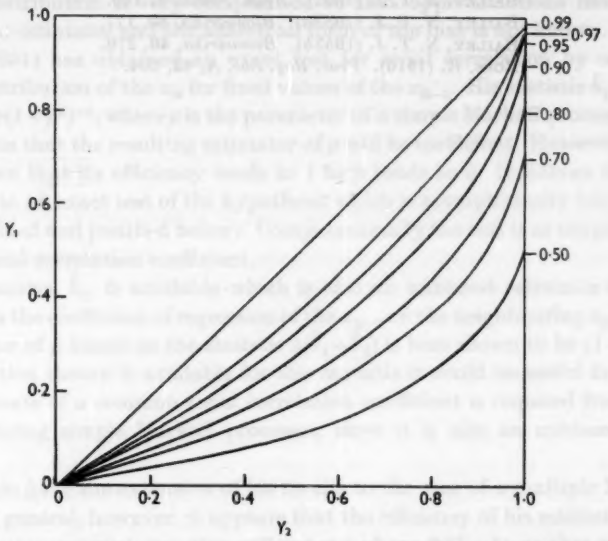


Fig. 1. (Y_1, Y_2) curves for $m = 2$ and $\gamma = 0{\cdot}2$ and different initial values: $Y_1(0) = 0{\cdot}99$, 0·97, 0·95, 0·90, 0·80, 0·70 and 0·50 and $Y_2(0) = 1$ in all cases.

of cases, for different numbers m of communities and different values of γ, and this last feature appears to be generally true and is particularly noticeable for small m (e.g. $m = 2$) and reasonably large values of the initial amount of infection in one community (e.g. above 10%) before infection in that community begins to affect susceptibles in the others. This calls to mind the remark of Ross (1916): 'It is obvious from the mere examination of many curves of epidemics that they are often remarkably symmetrical bell-shaped curves which, however, frequently tend to fall somewhat more slowly than they rise....'

A method of illustrating the solutions of the epidemic equations is shown in Fig. 1, where for different initial conditions (Y_1, Y_2) curves are plotted for the case $m = 2$ and $\gamma = 0{\cdot}2$. The curves shown are for $Y_1(0) = 0{\cdot}99$, 0·97, 0·95, 0·90, 0·80, 0·70 and 0·50, and $Y_2(0) = 1$ in all cases. Contours for the time parameter s, giving the times that points on the (Y_1, Y_2) curves are reached from the start of the epidemic, and contours showing the value of the total epidemic rate as the epidemic proceeds can also be put on this same graph. To avoid

confusion they have not been included in Fig. 1. but the latter are in fact conics. In this particular example. we see from equations (42) with $m = 2$. that the contours

$$-(dY_1/ds + dY_2/ds) = k$$

are concentric ellipses

$$Y_1^2 + 2\gamma Y_1 Y_2 + Y_2^2 - (1+\gamma)(Y_1 + Y_2) = k,$$

having their common centre at the point $Y_1 = Y_2 = \frac{1}{2}$, their minor axes along the line $Y_1 = Y_2$ of lengths $\sqrt{\{(1+\gamma+2k)/(2(1+\gamma))\}}$ and their major axes (perpendicular to $Y_1 = Y_2$) of lengths equal to $\sqrt{\{(1+\gamma)/(1-\gamma)\}}$ times the corresponding minor axes.

REFERENCES

BAILEY, N. T. J. (1952). *Appl. Statist.* **1**, 149.
BAILEY, N. T. J. (1953*a*). *Biometrika*, **40**, 177.
BAILEY, N. T. J. (1953*b*). *Biometrika*, **40**, 279.
ROSS, R. (1916). *Proc. Roy. Soc.* A, **92**, 204.

EXACT TESTS FOR SERIAL CORRELATION

By E. J. HANNAN

Australian National University, Canberra, A.C.T.

1. Introduction

In testing for first-order serial correlation in a stationary time series it is customary to use some form of the first sample serial correlation coefficient. Though this statistic will, presumably, be a most efficient estimator of the serial correlation ρ it has two disadvantages: (*a*) its exact distribution is very complicated so that approximations have to be used; (*b*) it is a biased estimator and the analytical form of the bias is not known.

Ogawara (1951) has obtained an exact test for serial correlation by considering the conditional distribution of the x_{2l} for fixed values of the x_{2l-1}. His statistic $\hat{b}_1$ is an unbiased estimator of $2\rho(1+\rho^2)^{-1}$, where ρ is the parameter of a simple Markoff process. At first sight it seems obvious that the resulting estimator of ρ will be inefficient. However, in this paper it will be shown that its efficiency tends to 1 as ρ tends to 0. It follows that Ogawara's statistic leads to an exact test of the hypothesis which is asymptotically fully efficient (this criterion is defined and justified below). Computationally the test is as simple as that based on the first serial correlation coefficient.

Another statistic, $\hat{b}_2$, is available which is also an unbiased estimator of $2\rho(1+\rho^2)^{-1}$. This statistic is the coefficient of regression of the x_{2l-1} on the neighbouring x_{2l}. The efficiency of the estimator of ρ based on the statistic $\frac{1}{2}(\hat{b}_1+\hat{b}_2)$ is here shown to be $(1-\rho^2)$. While no exact distribution theory is available for this statistic it could be useful in circumstances where an estimate of a common serial correlation coefficient is required from a number of otherwise differing simple Markoff processes, since it is also an unbiased estimator of $2\rho(1+\rho^2)^{-1}$.

Ogawara also gave the extension of his results to the case of a multiple Markoff process of order h. In general, however, it appears that the efficiency of his estimates of the coefficients of the autoregressive equation will not rise above $2/(h+1)$, so that only in the first-order case are they ever fully efficient.

The problem of testing for serial correlation in the residuals from a regression equation has been considered by Moran (1950) and Durbin & Watson (1950, 1951). Moran obtained asymptotic formulae for the mean and variance of the first circular serial correlation coefficient computed from the residuals from a least-squares regression on a single independent variate. The exact distribution of the test statistic used by Durbin & Watson is of a complicated form and is only known for regression vectors satisfying certain conditions. However, upper and lower bounds to the significance points, valid for any regression vectors, have been tabulated. In cases where the test statistic falls between the appropriate bounds an approximate test was suggested, based on the use of a beta distribution with the same mean and variance as the test statistic.

In this paper it will be shown that Ogawara's method can be extended to give an exact test, for the independence of the residuals from the regression equation, based on the z distribution. Again, a statistic is obtained which is an unbiased estimator of $2\rho(1+\rho^2)^{-1}$ in the case where the residuals follow a first-order Markoff scheme. The estimator is, as ρ

tends to zero, asymptotically as efficient an estimator of ρ as the first serial correlation coefficient of the residuals and Durbin & Watson's statistic. It again follows that the statistic leads to a test which is asymptotically fully efficient.

At the same time an exact test against given values of the regression coefficients is obtained in the case in which the residuals follow a simple Markoff process.

One disadvantage of these tests results from the necessity of computing a separate least-squares regression involving $2k+1$ regressors, where k is the number of 'independent' variates in the original regression equation. The estimates of the coefficients in the original equation, obtained from this regression, will, under some circumstances, have an asymptotically smaller variance than the estimates obtained by straightforward least squares. However, over a wide range of these conditions the estimates of the regression coefficients obtained from the first differences of the observations will have an even smaller asymptotic variance, and over this range at least the variances of these last estimates will almost certainly be smaller for small samples.

Finally, it should be emphasized that, though certain exact tests have been obtained, the powers of these tests have been judged from their asymptotic properties. For really small samples these tests may be far from optimal.

2. The criterion of asymptotic relative efficiency of tests

If t_1 and t_2 are statistics, computed from a sample of size n, their asymptotic efficiency as tests of an hypothesis specified by a parameter value θ_0 is defined as

$$E(t_1, t_2 \mid \theta_0) = \lim_{n \to \infty} \frac{\left\{\left[\frac{\partial}{\partial \theta} \mathscr{E}(t_1)\right]_{\theta=\theta_0}\right\}^2}{V(t_1 \mid \theta = \theta_0)} \frac{V(t_2 \mid \theta = \theta_0)}{\left\{\left[\frac{\partial}{\partial \theta} \mathscr{E}(t_2)\right]_{\theta=\theta_0}\right\}^2}.$$

Here $\mathscr{E}(t_j)$ and $V(t_j)$ are respectively the expected value and variance of t_j.

The justification for this criterion rests upon a theorem due to Pitman (see Stuart, 1954). Pitman has shown that, under certain regularity conditions, for t_1 and t_2 with limiting normal distributions and variances of order n^{-1}, the ratio of the sample sizes required for t_1 and t_2 to have the same power against alternative values of θ which differ from θ_0 by quantities of order $n^{-\frac{1}{2}}$ is in the limit given by $E(t_1, t_2 \mid \theta_0)$.

3. The test for serial correlation in a stationary autoregressive process

Ogawara considered the stochastic process $(x_t - m) = \rho(x_{t-1} - m) + \epsilon_t$, where ϵ_t is normally distributed with 0 mean and variance $\sigma^2(1-\rho^2)$, and m is the mean of the process. Here $|\rho| < 1$.

He showed that in the conditional distribution of the x_{2t} $(t = 1, \ldots, n)$ for fixed values of x_{2t-1} $(t = 1, \ldots, n+1)$, the parameter $b = 2\rho(1+\rho^2)^{-1}$ appears as the regression coefficient of the x_{2t} on the fixed variates $\frac{1}{2}(x_{2t-1} + x_{2t+1}) = x'_t$. Ogawara pointed out that the statistic

$$\hat{b}_1 = \frac{\sum_1^n (x'_t - \bar{x}')(x_{2t} - \bar{x}_2)}{\Sigma (x'_t - \bar{x}')^2}$$

is then the maximum-likelihood estimator of b, and is unbiased, and that

$$F = \frac{(\hat{b}_1 - b)^2 \sum_1^n (x_t' - \bar{x}')^2}{\sum_1^n (x_{2t} - \hat{a} - \hat{b}_1 x_t')^2} (n-2),$$

which has the F distribution with 1 and $n-2$ degrees of freedom, can be used as an exact test for any assigned value of b. Here $\hat{a} = \bar{x}_2 - \hat{b}_1 \bar{x}'$.

For the test of significance of the hypothesis $\rho = 0$ the statistic F reduces to

$$(n-2)\, r^2 (1-r^2)^{-1},$$

where r is the coefficient of correlation between x_{2t} and $\frac{1}{2}(x_{2t-1} + x_{2t+1})$. This test is, therefore, computationally as simple as that based on the first serial correlation coefficient.

The variance of $\hat{b}_1$, in the conditional distribution of the x_{2t}, is

$$\frac{\sigma^2(1-\rho^2)}{1+\rho^2} \frac{1}{\sum_1^n (x_t' - \bar{x}')^2}.$$

and since this converges in probability as n increases to

$$\frac{\sigma^2(1-\rho^2)\, 4}{(1+\rho^2)\, 2n\sigma^2(1+\rho^2)} = \frac{2}{n} \frac{(1-\rho^2)}{(1+\rho^2)^2},$$

the conditional variance of $\hat{b}_1$ also converges in probability to†

$$\frac{2}{n} \frac{1-\rho^2}{(1+\rho^2)^2}.$$

Given $\hat{b}_1$ we can estimate ρ by

$$\hat{\rho}_1 = \frac{1 - \sqrt{(1 - \hat{b}_1^2)}}{\hat{b}_1}.$$

This is not, of course, unbiased. Its variance will be, in the limit,

$$\frac{2}{n} \frac{1-\rho^2}{(1+\rho^2)^2} \frac{(1+\rho^2)^4}{4(1-\rho^2)^2} = \frac{1}{2n} \frac{(1+\rho^2)^2}{(1-\rho^2)}.$$

The variance of r_1, the first sample serial correlation coefficient, will tend to $\frac{1}{2}(1-\rho^2)/n$ (Bartlett, 1946), and since it seems that this will be asymptotically most efficient (Wald, 1948) the efficiency of Ogawara's estimate, $\hat{\rho}_1$, will be

$$\left(\frac{1-\rho^2}{1+\rho^2}\right)^2.$$

The efficiency of $\hat{\rho}_1$ for certain values of ρ is shown in Table 1.

Table 1

$\lvert\rho\rvert$	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
Efficiency of $\hat{\rho}_1$	0·96	0·85	0·70	0·52	0·36	0·22	0·12	0·05	0·01

† For any observed $\hat{b}_1$ the variance, for fixed x_{2t}, will be $\sigma^2(1-\rho^2)\left\{(1+\rho^2)\sum_1^n (x_t' - \bar{x}')^2\right\}^{-1}$ of course.

Another unbiased estimator of b is available in the statistic

$$\hat{b}_2 = \frac{\sum_1^{n-1} (x_t'' - \bar{x}'') x_{2t+1}}{\sum_1^{n-1} (x_t'' - \bar{x}'')^2},$$

where $x_t'' = \frac{1}{2}(x_{2t} + x_{2t+2})$.

The statistic $\hat{b} = \frac{1}{2}(\hat{b}_1 + \hat{b}_2)$ will also be unbiased and its variance will be

$$\tfrac{1}{4}\operatorname{var}(\hat{b}_1) + \tfrac{1}{4}\operatorname{var}(\hat{b}_2) + \tfrac{1}{2}\operatorname{cov}(\hat{b}_1 \hat{b}_2) = \frac{1-\rho^2}{n(1+\rho^2)^2} + \tfrac{1}{2}\operatorname{cov}(\hat{b}_1 \hat{b}_2).$$

To the order n^{-1} the sample means can be neglected and the covariance of $\hat{b}_1$ and $\hat{b}_2$ will be (to this order):

$$\frac{1}{(1+\rho^2)^2}\operatorname{cov}(a_1 a_2) + \frac{\rho^2}{(1+\rho^2)^4}\operatorname{cov}(c_1 c_2) - \frac{\rho}{(1+\rho^2)^3}\{\operatorname{cov}(a_1 c_2) + \operatorname{cov}(c_1 a_2)\},$$

where

$$a_1 = \frac{1}{\sigma^2 n}\sum_1^n x_{2t}(x_{2t-1} + x_{2t+1}), \quad a_2 = \frac{1}{\sigma^2 n}\sum_1^{n-1} x_{2t+1}(x_{2t} + x_{2t+2}),$$

$$c_1 = \frac{1}{\sigma^2 n}\sum_1^n (x_{2t-1} + x_{2t+1})^2, \quad c_2 = \frac{1}{\sigma^2 n}\sum_1^{n-1} (x_{2t} + x_{2t+2})^2.$$

By straightforward algebra we obtain, to the order n^{-1},

$$\operatorname{cov}(\hat{b}_1 \hat{b}_2) = \frac{1}{n}\left\{\frac{1}{(1+\rho^2)^2}\left[\frac{2+10\rho^2+6\rho^4-2\rho^6}{1-\rho^4}\right] + \frac{\rho^2}{(1+\rho^2)^4}\left[\frac{40\rho^2+40\rho^4-8\rho^6-8\rho^8}{1-\rho^4}\right] - \frac{2\rho}{(1+\rho^2)^3}\left[\frac{12\rho+20\rho^3+4\rho^5-4\rho^7}{1-\rho^4}\right]\right\}.$$

The asymptotic variance of $\hat{b}$, therefore, is

$$\operatorname{var}(\hat{b}_1)\left[\frac{1-\rho^2}{(1+\rho^2)^2}\right].$$

Hence, the variance of the estimator $\hat{\rho} = [1 - \sqrt{(1-\hat{b}^2)}]\hat{b}$ tends to $(2n)^{-1}$, so that its efficiency is $(1-\rho^2)$.

Table 2 shows the efficiency of this statistic for certain values of ρ.

Table 2

$\lvert\rho\rvert$	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9
Efficiency of $\hat{\rho}$	0·99	0·96	0·91	0·84	0·75	0·64	0·51	0·36	0·19

Though no exact distribution theory is available for this statistic, yet, in spite of its inefficiency, it could prove useful in circumstances where an estimate of a common serial correlation coefficient is required from a number of otherwise differing simple Markoff processes.† The lack of bias in $\frac{1}{2}(\hat{b}_1 + \hat{b}_2)$ would enable the several estimates to be combined

† An example (due to E. J. Williams) has arisen in connexion with the serial correlation of measurements at points along a log. The averaging of estimates from different logs, the only method of substantially increasing the number of observations, would not result in a consistent estimator if the estimates were obtained from the first sample serial correlation coefficient.

and an estimator of ρ to be obtained which would be consistent. The loss of efficiency, for moderate ρ, would not be large.

The statistics $\hat{b}_1$, $\hat{b}_2$ and $b = \frac{1}{2}(\hat{b}_1 + \hat{b}_2)$ have one major disadvantage however. The use of the formula

$$\hat{\rho}_j = [1 - \sqrt{(1 - \hat{b}_j^2)}]/\hat{b}_j$$

will lead to a consistent estimator of the first serial correlation coefficient only if the underlying process is truly a simple Markoff process.

The fact that the efficiency of $\hat{\rho}_j$ tends to 1 as ρ tends to 0 is important, however, for this shows that the test of significance based on $\hat{b}_j$ $(j = 1, 2)$ will be asymptotically fully efficient. Though nothing certain can be said, from this result, about the power of the test in small samples it seems likely that it will be nearly as powerful as that based on r_1 against the range of alternatives represented by simple Markoff processes. The test based on $\hat{b}_j$ has the advantage of being exact.

Ogawara extended his results to a multiple Markoff process of order h. In this case the regression of the variate $x_{t(h+1)}$ on the variates $x'_{pt} = \frac{1}{2}(x_{t(h+1)+p} + x_{t(h+1)-p})$ for $p = 1, \ldots, h$ is considered. In the first-order case the basic test statistic is the correlation coefficient between x_{2t} and $\frac{1}{2}(x_{2t-1} + x_{2t+1})$, and the numerator of this statistic (apart from the factor $\frac{1}{2}$) differs only in the corrections for the means from the numerator of the first sample serial correlation coefficient. The corresponding test statistics in the general case are the total and partial correlation coefficients between $x_{t(h+1)}$ and the x'_{pt}. Now, however, certain cross-products are omitted as compared with the total and partial sample serial correlation coefficients so that the power of the tests is reduced. It seems that the efficiency of the estimates of the coefficients of the autoregressive equation (from which the asymptotic power of the tests can be gauged) will be at a maximum when all of the coefficients are zero. In this case the efficiency is $2(h+1)^{-1}$. In general, it will be lower. For example, the efficiency of the estimate of a_2, in the process $x_t + a_1 x_{t-1} + a_2 x_{t-2} = \epsilon_t$ in case $a_2 = 0$, is $\frac{2}{3}(1 + a_1^2)^{-1}$. It appears, therefore, that although exact tests for partial serial correlation are obtained by Ogawara's approach these tests will be less powerful than those provided by the usual serial coefficients.

4. THE TEST FOR SERIAL CORRELATION IN THE RESIDUALS FROM A REGRESSION EQUATION

Consider the regression $\quad y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + \ldots + \beta_k x_{kt} + \epsilon_t,$

where

(1) $\epsilon_t = \rho \epsilon_{t-1} + \eta_t$ and η_t is $N(0, \sigma(1-\rho^2)^{\frac{1}{2}})$; $|\rho| < 1$

(2) The ϵ_t are independent of the x_{jt}.

Then the conditional distribution of the y_{2t} for fixed $x_{j,t}$ and fixed ϵ_{2t+1} $(t = 0, \ldots, n)$ is

$$(2\pi\sigma_0^2)^{-\frac{1}{2}n} \exp -\frac{1}{2\sigma_0^2} \sum_{t=1}^{n} \left(y_{2t} - \gamma_0 - \sum_{j=1}^{2k+1} \gamma_j z_{jt} \right)^2 \prod_{1}^{n} dy_{2t}.$$

Here

$$\sigma_0^2 = \frac{\sigma^2(1-\rho^2)}{1+\rho^2}, \quad \gamma_0 = \alpha(1 - \gamma_{k+1}),$$

$$\gamma_j = \beta_j \quad \text{and} \quad z_{j,t} = x_{j,2t} \quad (j = 1, \ldots, k),$$

$$\gamma_{k+1} = \frac{2\rho}{1+\rho^2} \quad \text{and} \quad z_{k+1,t} = \tfrac{1}{2}(y_{2t-1} + y_{2t+1}),$$

$$\gamma_j = -\gamma_{k+1}\beta_j \quad \text{and} \quad z_{j,t} = \tfrac{1}{2}(x_{j,2t-1} + x_{j,2t+1}) \quad (j = k+2, \ldots, 2k+1).$$

If the information involved in the knowledge that $\gamma_{j+k+1} = -\gamma_{k+1}\gamma_j$ $(j = 1, \ldots, k)$ is discarded and the coefficients γ_j are estimated by the straightforward least-squares procedure, these estimates will have the usual properties of least-squares estimates in the classic case. They will not be maximum-likelihood estimates. However, the computation of the maximum-likelihood estimates will involve the solution of a system of second-order equations and no exact tests will be available.

Representing the least-squares estimates of the γ_j by $\hat{\gamma}_j$, the hypothesis $\rho = 0$ can be tested by the use of the statistic

$$F_{1,n-2k-2} = (n-2k-2)\frac{\hat{\gamma}_{k+1}^2}{\sigma^{*2}}\frac{|\mathbf{L}|}{|\mathbf{L}_{k+1,k+1}|},$$

which has the F distribution with 1 and $n-2k-2$ degrees of freedom. Here σ^{*2} is the variance of the residuals from the least-squares regression of y_{2t} on the z_{jt} while $\mathbf{L}$ is the covariance matrix of the z_{jt}: $|\mathbf{L}_{k+1,k+1}|$ is the cofactor of the element in the indicated row and column. The test is, of course, equivalent to a test of the partial correlation of y_{2t} and $\frac{1}{2}(y_{2t-1}+y_{2t+1})$ with the effects of the z_{jt} $(j \neq k+1)$ removed.

Tests of significance and confidence limits for the parameters β_j can be obtained from statistics such as

$$F_{1,n-2k-2} = (n-2k-2)\left\{(\hat{\gamma}_j-\beta_j)^2\frac{|\mathbf{L}|}{|\mathbf{L}_{jj}|}\right\}\sigma^{*-2} \quad (j = 1, \ldots, k),$$

which have the F distribution with 1 and $(n-2k-2)$ degrees of freedom. Similarly, the multiple correlation of y and the x_j may be tested by the statistic

$$F_{k,n-2k-2} = \frac{(n-2k-2)}{k}\left\{\sum_{ij=1}^{k}\hat{\gamma}_i\hat{\gamma}_j\frac{|\mathbf{\Lambda}_{ij}|}{|\mathbf{\Lambda}|}\right\}\sigma^{*-2}.$$

Here $\mathbf{\Lambda}$ is the matrix formed from the first k rows and columns of $\mathbf{L}^{-1}$.

These last tests are exact for any value of ρ, though the power of the tests will depend upon that parameter.

In order to examine the limiting variances and covariances of the estimates of the γ_j it is necessary to obtain the stochastic limit of the matrix $\mathbf{L}$.

If the vector $\{z_{jt}\}$ is written $\mathbf{Z}_t = \{\mathbf{Z}_{1t}, z_{k+1,t}, \mathbf{Z}_{2t}\}$ and the corresponding vector of sample means $\overline{\mathbf{Z}} = \{\overline{\mathbf{Z}}_1, \bar{z}_{k+1}, \overline{\mathbf{Z}}_2\}$, then the covariance matrix $\mathbf{L}$ is

$$\mathbf{L} = \frac{1}{n}\sum_{t=1}^{n}\begin{bmatrix} (\mathbf{Z}_{1t}-\overline{\mathbf{Z}}_1)\mathbf{Z}_{1t}' & (\mathbf{Z}_{1t}-\overline{\mathbf{Z}}_1)z_{k+1,t} & (\mathbf{Z}_{1t}-\overline{\mathbf{Z}}_1)\mathbf{Z}_{2t}' \\ (z_{k+1,t}-\bar{z}_{k+1})\mathbf{Z}_{1t}' & (z_{k+1,t}-\bar{z}_{k+1})z_{k+1,t} & (z_{k+1,t}-\bar{z}_{k+1})\mathbf{Z}_{2t}' \\ (\mathbf{Z}_{2t}-\overline{\mathbf{Z}}_2)\mathbf{Z}_{1t}' & (\mathbf{Z}_{2t}-\overline{\mathbf{Z}}_2)z_{k+1,t} & (\mathbf{Z}_{2t}-\overline{\mathbf{Z}}_2)\mathbf{Z}_{2t}' \end{bmatrix}.$$

Since $z_{k+1,t} = \alpha + \boldsymbol{\beta}'\mathbf{Z}_{2t} + \frac{1}{2}(\epsilon_{2t-1}+\epsilon_{2t+1})$, where $\boldsymbol{\beta}$ is the vector of regression coefficients, under fairly weak restrictions on the nature of the processes generating the x_{jt}, this matrix will have a limit in probability of the form

$$\begin{bmatrix} \mathbf{X}_0 & \mathbf{X}_1\boldsymbol{\beta} & \mathbf{X}_1 \\ \boldsymbol{\beta}'\mathbf{X}_1' & \frac{1}{2}\sigma^2(1+\rho^2)+\boldsymbol{\beta}'\mathbf{X}_2\boldsymbol{\beta} & \boldsymbol{\beta}'\mathbf{X}_2 \\ \mathbf{X}_1' & \mathbf{X}_2\boldsymbol{\beta} & \mathbf{X}_2 \end{bmatrix},$$

where $\mathbf{X}_0$, $\mathbf{X}_1$ and its transpose and $\mathbf{X}_2$ are the stochastic limits of the matrix in the corners of $\mathbf{L}$.

Since the determinantal value of $\mathbf{L}$ will be almost everywhere different from 0 (if the joint distribution of the x_j is not singular) the matrix $\mathbf{L}^{-1}$ will converge in probability to the inverse of the matrix just written.

Since the covariance matrix of the $\hat{\gamma}_j$ (for $j > 0$) is $\sigma_0^2(nL)^{-1}$, it is readily seen that the variance of $\hat{\gamma}_{k+1}$ converges in probability to

$$\frac{2}{n}\frac{1-\rho^2}{(1+\rho^2)^2}.$$

Hence the variance of the estimator of ρ obtained from this statistic is

$$\frac{1}{2n}\frac{(1+\rho^2)^2}{(1-\rho^2)},$$

as in the case of the estimation of the parameter of a simple Markoff process by Ogawara's method.

The test statistic, d, used by Durbin & Watson, is asymptotically equal to $2(1-r_1)$, where r_1 is the first serial correlation coefficient computed from the actual residuals from a least-squares regression. It can be seen from the formulae (5), (6), (7) and (8) given on p. 164 of their paper (1951) that the variance of this statistic (when $\rho = 0$) will be, asymptotically, $2n^{-1}$ when computed from $2n$ observations. It follows that the estimator of ρ obtained from $\hat{\gamma}_{k+1}$ is as efficient as that gained from the statistic d, when $\rho = 0$. The asymptotic relative efficiency of d and $\hat{\gamma}_{k+1}$ against the hypothesis $\rho = 0$ is therefore unity. It is easy to show that the maximum-likelihood estimator of ρ will have an asymptotic variance $(2n)^{-1}$ (from $2n$ observations) when ρ equals zero. It therefore appears that d and $\hat{\gamma}_{k+1}$ lead to tests of the hypothesis, $\rho = 0$, which are asymptotically fully efficient.

The covariance matrix of the $\hat{\gamma}_j$ $(j = 1, \ldots, k)$ converges in probability to

$$\left(\frac{\sigma^2}{n}\frac{1-\rho^2}{1+\rho^2}\right)\mathbf{X},$$

where $\mathbf{X}$ is the matrix in the top left-hand corner of the inverse of

$$\begin{bmatrix} \mathbf{X}_0 & \mathbf{X}_1 \\ \mathbf{X}_1' & \mathbf{X}_2 \end{bmatrix}.$$

If the x_j are serially independent the covariance matrix of the $\hat{\gamma}_j$ $(j = 1, \ldots, k)$ becomes in the limit

$$\frac{\sigma^2}{n}\frac{1-\rho^2}{1+\rho^2}\mathbf{X}_0^{-1}.$$

But these $\hat{\gamma}_j$ are unbiased estimators of the β_j. The covariance matrix of the straightforward least-squares estimates of the β_j is, in this very special case (Wold, 1953, p. 213),

$$\frac{\sigma^2}{2n}\mathbf{X}_0^{-1}.$$

The relative efficiency of the estimates of the β_j by the two methods is then $\frac{1}{2}(1+\rho^2)(1-\rho^2)^{-1}$. The estimates obtained from the $\hat{\gamma}_j$ will, in this case, be more efficient than those from least squares when $\rho\sqrt{3} > 1$.

In general, the relative efficiency of the two methods will depend upon the correlogram of the x_{jt} as well as ρ. For example, in the case where there is only one regressor, which follows a simple Markoff process with parameter ρ_1, the variance of $\hat{\gamma}_1$ will be, in the limit,

$$\text{var}\,(\hat{\gamma}_1) = \frac{\sigma^2}{n}\left(\frac{1-\rho^2}{1+\rho^2}\right)\left(\frac{1+\rho_1^2}{1-\rho_1^2}\right)\frac{1}{\sigma_1^2},$$

where σ_1^2 is the variance of x_1.

The variance of the straightforward least-squares estimate from $2n$ observations will then tend to (Wold, 1953, p. 211)

$$\text{var}\,(\hat{\beta}_1) = \frac{\sigma^2}{2n}\frac{1+\rho\rho_1}{1-\rho\rho_1}\frac{1}{\sigma_1^2}.$$

Table 3. $\text{var}\,(\hat{\beta}_1)/\text{var}\,(\hat{\gamma}_1)$

ρ / ρ_1	−0·8	−0·6	−0·4	−0·2	0	0·2	0·4	0·6	0·8
0	2·28	1·06	0·69	0·54	0·50	0·54	0·69	1·06	2·28
0·2	1·52	0·77	0·54	0·46	0·46	0·54	0·75	1·25	2·90
0·4	0·85	0·47	0·36	0·33	0·36	0·42	0·69	1·26	3·20
0·6	0·38	0·24	0·20	0·20	0·24	0·32	0·63	1·06	3·05
0·8	0·11	0·08	0·08	0·09	0·11	0·16	0·29	0·66	2·28

The relative efficiency $\text{var}\,(\hat{\beta}_1)/\text{var}\,(\hat{\gamma}_1)$ is shown in Table 3. This table will also give the relative efficiencies of the $\hat{\gamma}_j$ $(j = 1, \ldots, k)$ for a k variate regression when each of the x_j follows a simple Markoff process with the same parameter ρ_1.

The limiting variance of $\hat{\gamma}_0$ is $\dfrac{\sigma^2}{n}\dfrac{1-\rho^2}{1+\rho^2}$.

The corresponding estimate of α is $\hat{\gamma}_0(1-\hat{\gamma}_{k+1})^{-1}$, and this has the asymptotic variance

$$\frac{\sigma^2}{n}\frac{1+\rho}{(1-\rho)^3}\left(1+\rho^2+\frac{2\alpha^2}{\sigma^2}\right).$$

The variance of the straightforward least-squares estimate of α will tend to

$$\frac{\sigma^2}{n}\frac{1+\rho}{1-\rho},$$

so that the estimate obtained from $\hat{\gamma}_0$ is always relatively inefficient and is very inefficient for values of ρ near 1. In such cases a better estimate than that obtained from $\hat{\gamma}_0$ would be $\hat{\alpha} = \bar{y} - \sum_1^k \hat{\gamma}_j \bar{x}_j$, where the means are obtained from all of the observations.

A common procedure, when positive serial correlation is suspected in the residuals, is to estimate the regression coefficients from the first differences of the y_t and x_{jt} (see Cochrane & Orcutt, 1949). Again the efficiency of this method will depend upon the correlogram of the x_{jt}. When there is only one regressor, which follows a simple Markoff process with

variance σ_1^2 and serial correlation ρ_1, the asymptotic variance of the least-squares estimate of β_1, from the first differences, can be shown to be (when estimated from $2n$ observations)

$$\frac{\sigma^2}{\sigma_1^2}\frac{1}{4n}\frac{(1-\rho)}{(1-\rho_1)}\left\{\frac{3-\rho\rho_1-\rho-\rho_1}{1-\rho\rho_1}\right\}.$$

If $\rho \geqslant 0$ this will be smaller than the variance of the estimator $\hat{\gamma}_1$.

It must be emphasized that the validity of the tests of the regression coefficients, together with the estimates of these coefficients (including the estimate of ρ derived from $\hat{\gamma}_{k+1}$), depends on the specification of the errors in the regression equation as a simple Markoff process being correct. The test of significance of ρ is, of course, always an exact test.

Example. Durbin & Watson (1951) use economic data due to A. R. Prest (1949) to demonstrate their method. The original data, given in their paper, cover the period 1870–1938, and show the logarithm of the consumption of spirits per head in the United Kingdom (y_t), the logarithm of real income per head (x_{1t}) and the logarithm of the relative price of spirits (x_{2t}). The estimates of the coefficients from the least-squares regression of y on x_1 and x_2 are

$$\hat{\beta}_1 = -0{\cdot}120, \quad \hat{\beta}_2 = -1{\cdot}228.$$

Their statistic $d = 0{\cdot}2488$ and is far below the lower bound to the significance point for d at the 1 % level, indicating a highly significant positive serial correlation.

We will apply the method presented in this paper to this example.

The coefficients $\hat{\gamma}_j$ are given by

$$\begin{bmatrix}\hat{\gamma}_1\\ \hat{\gamma}_2\\ \hat{\gamma}_3\\ \hat{\gamma}_4\\ \hat{\gamma}_5\end{bmatrix} = \begin{bmatrix} 0{\cdot}307 & 0{\cdot}490 & -0{\cdot}651 & 0{\cdot}295 & 0{\cdot}495\\ 0{\cdot}490 & 1{\cdot}445 & -1{\cdot}818 & 0{\cdot}474 & 1{\cdot}423\\ -0{\cdot}651 & -1{\cdot}818 & 2{\cdot}423 & -0{\cdot}631 & -1{\cdot}815\\ 0{\cdot}295 & 0{\cdot}474 & -0{\cdot}631 & 0{\cdot}288 & 0{\cdot}478\\ 0{\cdot}495 & 1{\cdot}423 & -1{\cdot}815 & 0{\cdot}478 & 1{\cdot}417\end{bmatrix}^{-1}\begin{bmatrix}-0{\cdot}636\\ -1{\cdot}818\\ 2{\cdot}401\\ -0{\cdot}621\\ -1{\cdot}804\end{bmatrix} = \begin{bmatrix}0{\cdot}795\\ -0{\cdot}693\\ 0{\cdot}950\\ -0{\cdot}792\\ 0{\cdot}630\end{bmatrix}$$

$$\hat{\gamma}_0 = 0{\cdot}218, \quad \sigma^{*2} = 0{\cdot}000232.$$

The elements of the 5×5 matrix are $\sum_1^{34}(z_{it}-\bar{z}_i)(z_{jt}-\bar{z}_j)$. For example,

$$0{\cdot}495 = \Sigma\tfrac{1}{2}(x_{2,2t-1}+x_{2,2t+1}-\bar{x}_2')(x_{1,2t}-\bar{x}_{1,2}),$$

where

$$\bar{x}_2' = \tfrac{1}{34}\Sigma\tfrac{1}{2}(x_{2,2t-1}+x_{2,2t+1}),$$

$$\bar{x}_{1,2} = \tfrac{1}{34}\Sigma x_{1,2t}.$$

All of the parameters $\hat{\gamma}_j$ are significant at the 1 % point. The value of F corresponding to $\hat{\gamma}_3$ equals 308 in fact. However, the 5 % confidence intervals for γ_1 and γ_2 do not include the straightforward least-squares estimates. They are:

$$0{\cdot}339 \leqslant \gamma_1 \leqslant 1{\cdot}251, \quad -0{\cdot}991 \leqslant \gamma_2 \leqslant -0{\cdot}405.$$

On prior grounds the value of $-0{\cdot}120$ for $\hat{\beta}_1$ is unacceptable, for it is very unlikely that the demand for spirits would fall as income rises. In the original work of A. R. Prest (1949) the coefficients were estimated from the first differences of the observations and the resulting regression coefficients were

$$\hat{\beta}_1 = 0{\cdot}736, \quad \hat{\beta}_2 = -0{\cdot}861,$$

both of which lie within the 5 % confidence intervals for γ_1 and γ_2.

The variates x_1 and x_2 have high positive serial correlation coefficients of about the same value. Together with the high positive serial correlation of the residuals this suggests that the extension of Ogawara's method here presented will give estimators of β_1 and β_2 which will be at least as efficient as those from the straightforward least-squares procedure. This is to some extent borne out by the results.

I wish to thank Prof. P. A. P. Moran for suggesting this subject and its developments to me and for his advice in the research done and in the preparation of this paper. I should also like to thank the referee for a number of suggestions and, in particular, for pointing out the importance of the first difference transformation.

REFERENCES

BARTLETT, M. S. (1946). On the theoretical specification and sampling properties of auto-correlated time series. *J.R. Statist. Soc. Suppl.* **8**, 27.

COCHRANE, D. & ORCUTT, G. H. (1949). Application of least squares regression to relationships containing auto-correlated error terms. *J. Amer. Statist. Ass.* **44**, 32.

DURBIN, J. & WATSON, G. S. (1950, 1951). Testing for serial correlation in least squares regression. I. *Biometrika*, **37**, 409; II. *Biometrika*, **38**, 159.

MORAN, P. A. P. (1950). A test for serial independence of residuals. *Biometrika*, **37**, 178.

OGAWARA, M. (1951). A note on the test of serial correlation coefficients. *Ann. Math. Statist.* **22**, 115.

PREST, A. R. (1949). Some experiments in demand analysis. *Rev. Econ. Statist.* **31**, 33.

STUART, A. (1954). Asymptotic relative efficiencies of distribution-free tests of randomness against normal alternatives. *J. Amer. Statist. Ass.* **49**, 147.

WALD, A. (1948). Asymptotic properties of the maximum likelihood estimate of the unknown parameter of a discrete stochastic process. *Ann. Math. Statist.* **19**, 40.

WOLD, H. (1953). *Demand Analysis*. New York: John Wiley and Sons.

ON THE EFFICIENCY OF PROCEDURES FOR SMOOTHING PERIODOGRAMS FROM TIME SERIES WITH CONTINUOUS SPECTRA

By M. S. BARTLETT and J. MEDHI
University of Manchester

1. Introduction

Let X_r $(r = 1, 2, \ldots, n)$ be a set of consecutive observations from a real stochastic process X_t in discrete time. The process is assumed stationary, at least up to the second order, so that the stochastic average $E\{X_t\}$ is a constant (m), which for convenience is put zero, and the autocovariance function

$$E\{X_t X_{t+s}\} = w_s = \sigma_X^2 \rho_s \tag{1·1}$$

is a function only of the interval s. Then an autocorrelation function ρ_s $(s = 0, \pm 1, \pm 2, \ldots)$ is a valid one for some stationary process X_t if and only if

$$\rho_s = \int_0^\pi \cos(s\lambda)\, dF(\lambda), \tag{1·2}$$

where $F(\lambda)$, the (integrated) spectrum of the process, has the form of a distribution function defined in $(0, \pi)$ (Wold, 1938). Ignoring the possibility of a third singular component, we have further the relation

$$F(\lambda) = c_1^2 F_1(\lambda) + c_2^2 F_2(\lambda), \tag{1·3}$$

where $F_1(\lambda)$ is a step function, the (integrated) discrete spectrum, and $F_2(\lambda)$ has a non-zero derivative $f_2(\lambda)$, the continuous spectrum. In particular when c_1 is zero (as will now be assumed)

$$\rho_s = \int_0^\pi \cos(s\lambda) f(\lambda)\, d\lambda, \tag{1·4}$$

which has the inversion formula

$$f(\lambda) = \frac{1}{\pi} \sum_{s=-\infty}^{\infty} \rho_s \cos(s\lambda). \tag{1·5}$$

We define

$$g(\lambda) = 2\pi\sigma_X^2 f(\lambda),$$

so that

$$g(\lambda) = 2 \sum_{s=-\infty}^{\infty} w_s \cos(s\lambda). \tag{1·6}$$

The problem of estimating directly the spectral density function $f(\lambda)$ (or $g(\lambda)$), previously considered by Daniell (see the discussion to the paper by Bartlett, 1946), Bartlett (1950) and Grenander (1951), will be discussed further in this paper. Grenander & Rosenblatt (1952, 1954) have recently investigated also the problem of constructing an entire confidence band for the integrated density function $g(\lambda)$; for further suggestions on this problem, which will not be considered here, see Bartlett (1954).

2. Estimates of the spectral density function

The sample autocovariance, which will be defined as

$$C_s = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{r=1}^{n-s} X_r X_{r+s} & (|s| < n), \\ 0 & (|s| \geqslant n), \end{cases} \tag{2·1}$$

is, under wide conditions, a consistent estimate of w_s. If now we substitute C_s for w_s in (1·6), we obtain

$$2\sum_{s=-n}^{n} C_s \cos(s\lambda) = \frac{2}{n}\sum_{\mu,\nu=1}^{n} X_\mu X_\nu \cos[(\mu-\nu)\lambda]$$
$$= I(\lambda), \tag{2·2}$$

where $I(\lambda)$ is the periodogram intensity $G(\lambda)\,G^*(\lambda) \equiv A^2(\lambda) + B^2(\lambda)$, $G(\lambda)$ being calculated as

$$\sqrt{\frac{2}{n}}\sum_{r=1}^{n} X_r e^{i\lambda r} \tag{2·3}$$

(usually for integral p, where $\lambda = 2\pi p/n$). However, in spite of the relation (2·2), it is now known that $I(\lambda)$ does not provide a consistent estimate of $g(\lambda)$, owing to its large sampling fluctuations which do not diminish as n increases. The relevant stochastic properties of $I(\lambda)$ have been discussed by Slutsky (1927), Bartlett (1950, 1954), Grenander (1951) and others, and will merely be quoted when necessary. The further condition imposed on X_r for these properties to hold is that, if not normal, at least it has the linear form

$$X_t = \sum_{u=0}^{\infty} b_u Y_{t-u}, \tag{2·4}$$

where Y_t is a sequence of *independent* quantities with zero mean and constant variance σ^2.

In view of the sampling properties of $I(\lambda)$, 'smoothed' estimates of $g(\lambda)$ have been proposed. Thus Daniell put forward the estimate

$$g_D(\lambda) = \frac{1}{2h}\int_{\lambda-h}^{\lambda+h} I(\omega)\,d\omega; \tag{2·5}$$

and Bartlett's formula (cf. also Tukey and Hamming, 1949) for smoothing the periodogram $I(\lambda)$ is

$$g_B(\lambda) = 2\sum_{s=-n'+1}^{n'-1}\left(1-\frac{|s|}{n'}\right) C'_s \cos(s,\lambda), \tag{2·6}$$

where $C'_s = C_s \Big/ \left(1-\frac{|s|}{n}\right)$. The resolving power of this last procedure increases with n', and the smoothing with $m = n/n'$ (where n is the total length of the series), fluctuations being of order $1/\sqrt{m}$. Grenander subsequently examined a whole class of estimates containing the above two, by introducing a general weighting factor, $u_s(\lambda)$, in the formula

$$g_G(\lambda) = 2\sum_{s=-n}^{n} u_s(\lambda)\, C_s \cos(s\lambda), \tag{2·7}$$

or in alternative form (for suitable conditions on $u_s(\lambda)$)

$$g_G(\lambda) = \int_0^{\pi} I(\omega)\, w_\lambda(\omega)\,d\omega, \tag{2·8}$$

where the positive weighting function $w_\lambda(\omega)$ is expressible in terms of $u_s(\lambda)$ by the formula

$$w_\lambda(\omega) = \frac{1}{\pi}\sum_{s=-n}^{n} u_s(\lambda)\cos(s\lambda)\cos(s\omega). \tag{2·9}$$

The estimate $g_G(\lambda)$ has the asymptotic sampling properties

$$E\{g_G(\lambda)\} \sim \int_0^{\pi} g(\omega)\, w_\lambda(\omega)\,d\omega, \tag{2·10}$$

$$\operatorname{var}\{g_G(\lambda)\} \sim \frac{2\pi}{n}\int_0^{\pi} g^2(\omega)\, w_\lambda^2(\omega)\,d\omega. \tag{2·11}$$

The relation of Daniell's estimate $g_D(\lambda)$ in (2·5) to the general formula (2·8) is obvious; the other estimate in (2·6) has for large n the effective weighting factor

$$u_s(\lambda) = \begin{cases} 1 - \dfrac{|s|}{n'} & (|s| < n'), \\ 0 & (|s| \geqslant n'), \end{cases} \tag{2·12}$$

and the weighting function corresponding to $u_s(\lambda)$ in (2·12) is

$$w_\lambda(\omega) = \frac{1}{2\pi n'}\left[\frac{\sin^2\{\frac{1}{2}n'(\omega-\lambda)\}}{\sin^2\{\frac{1}{2}(\omega-\lambda)\}} + \frac{\sin^2\{\frac{1}{2}n'(\omega+\lambda)\}}{\sin^2\{\frac{1}{2}(\omega+\lambda)\}}\right]. \tag{2·13}$$

Clearly

$$E\{g_D(\lambda)\} \sim \frac{1}{2h}\int_{\lambda-h}^{\lambda+h} g(\omega)\,d\omega$$

$$\sim g(\lambda) \quad \text{for small } h \text{ only}, \tag{2·14}$$

and

$$\operatorname{var}\{g_D(\lambda)\} \sim \frac{\pi}{nh} g^2(\lambda) \quad \text{for small } h. \tag{2·15}$$

Grenander has given also the results

$$E\{g_B(\lambda)\} \sim g(\lambda) \tag{2·16}$$

and

$$\operatorname{var}\{g_B(\lambda)\} \sim \begin{cases} \dfrac{2}{3m} g^2(\lambda) & (\lambda \neq 0), \\ \dfrac{4}{3m} g^2(0) & (\lambda = 0). \end{cases} \tag{2·17}$$

3. Uncertainty relations

When Grenander investigated further the sampling properties of $g_G(\lambda)$, he imposed the condition

$$\int_0^\pi w_\lambda(\omega)\,d\omega = 1 \tag{3·1}$$

as a 'sort of asymptotic unbiasedness', but this condition (which is satisfied by the weighting function for $g_D(\lambda)$, even for h not small) appears unsatisfactory, and we shall use the correct condition for asymptotic unbiasedness from formula (2·10), viz.

$$\int_0^\pi g(w)\, w_\lambda(\omega)\,d\omega \sim g(\lambda). \tag{3·2}$$

With this condition the formula for $\operatorname{var}\{g_G(\lambda)\}$ in (2·11) yields

$$\frac{n}{2\pi}\operatorname{var}\{g_G(\lambda)\} \sim \int_0^\pi g^2(\omega)\, w_\lambda^2(\omega)\,d\omega \geqslant \frac{1}{\pi}\left[\int_0^\pi g(\omega)\, w_\lambda(\omega)\,d\omega\right]^2$$

$$\sim \frac{1}{\pi} g^2(\lambda),$$

whence asymptotically

$$\operatorname{var}\{g_G(\lambda)\} \geqslant \frac{2}{n} g^2(\lambda). \tag{3·3}$$

Moreover, the condition for equality is reached if

$$w_\lambda(\omega) = g(\lambda)/[\pi g(\omega)]. \tag{3·4}$$

This 'ideal' estimate thus has weighting function inversely proportional to $g(\omega)$,† not $g^2(\omega)$, as Grenander obtained with condition (3·1). However, it is still entirely theoretical in character, for it assumes a complete knowledge of the unknown spectral function $g(\omega)$. Furthermore, for different values of λ, the weighting function multiplying the stochastic quantity $I(\omega)$ remains inversely proportional to $g(\omega)$, and all the corresponding estimates $\hat{g}_T(\lambda)$ are stochastically equivalent.

In an attempt to overcome such difficultues, Grenander made use of a measure of 'resolvability', defined in general by the 'standard deviation' of ω for the weighting function $w_\lambda(\omega)$. As $w_\lambda(\omega)$ no longer necessarily satisfies the 'normalization' condition (3·1), this measure is modified here to

$$U_1 = \left[\int_0^\pi (\omega-\lambda)^2 w_\lambda(\lambda)\,d\omega \Big/ \int_0^\pi w_\lambda(\omega)\,d\omega\right]^{\frac{1}{2}}, \tag{3·5}$$

(note also that the deviation $w-\lambda$ is used, whether or not the 'mean' ω is exactly λ). It is further generalized to

$$U_1^{(r)} = \left[\int_0^\pi |\,\omega-\lambda\,|^r w_\lambda(\omega)\,d\omega \Big/ \int_0^\pi w_\lambda(\omega)\,d\omega\right]^{1/r} \quad (r>0), \tag{3·6}$$

mainly in order to stress the rather arbitrary character of such measures. In terms of $U_1^{(r)}$ and

$$U_2 \equiv \int_0^\pi g^2(\omega)\,w_\lambda^2(\omega)\,d\omega \sim \frac{n}{2\pi}\operatorname{var}\{g_G(\lambda)\}, \tag{3·7}$$

we may now generalize Grenander's 'uncertainty principle', connecting $U_1^{(r)}$ and U_2 by the following modification of his argument.

From the Tchebychev inequality

$$Q\{|\,\omega-\lambda\,| < 2U_1^{(r)}\} > \left(1-\frac{1}{2^r}\right),$$

where Q denotes 'probability' in terms of the formal frequency function

$$q(\omega) \equiv w_\lambda(\omega) \Big/ \int_0^\pi w_\lambda(\omega)\,d\omega,$$

$$\int_0^\pi q^2(\omega)\,d\omega \geqslant \int_{\lambda-2U_1^{(r)}}^{\lambda+2U_1^{(r)}} q^2(\omega)\,d\omega \geqslant \frac{1}{4U_1^{(r)}}\left[\int_{\lambda-2U_1^{(r)}}^{\lambda+2U_1^{(r)}} q(\omega)\,d\omega\right]^2$$

$$> \frac{1}{4U_1^{(r)}}\left[1-\frac{1}{2^r}\right]^2.$$

Let A, B be the upper and lower bounds of $g(\omega)$ in $(0,\pi)$, these being assumed (respectively) finite and non-zero. Then

$$U_2 \geqslant B^2\int_0^\pi w_\lambda^2(\omega)\,d\omega$$

$$> \frac{B^2 k_0^2}{4U_1^{(r)}}\left[1-\frac{1}{2^r}\right]^2,$$

where

$$k_r \equiv \int_0^\pi \omega^r w_\lambda(\omega)\,d\omega. \tag{3·8}$$

† The ideal weighting of $I(\omega)$ by a function varying as $1/g(\omega)$ is in entire accord with the asymptotic transformation to a uniform spectrum often desirable in spectral analysis (see, for example, Bartlett, 1954).

Now
$$Ak_0 \geqslant \int_0^\pi w_\lambda(\omega)\, g(\omega)\, d\omega \sim g(\lambda),$$

whence finally
$$U \equiv U_1^{(r)} U_2 > \tfrac{1}{4} g^2(\lambda) \left\{ \frac{B^2}{A^2} \left[1 - \frac{1}{2^r} \right]^2 \right\} > 0, \tag{3·9}$$

which shows that the product U of the two quantities $U_1^{(r)}$ and U_2 has a positive minimum.

The condition of asymptotic unbiasedness which we have imposed could perhaps be waived by considering the mean-square deviation of $g_G(\lambda)$ from $g(\lambda)$; this would, however, no longer give a criterion like U_2 asymptotically independent of n, so that it seems more convenient to adopt the conditions above.

4. Asymptotic efficiencies of $g_D(\lambda)$ and $g_B(\lambda)$

From (3·9) we may define a quantity
$$W \equiv U/g^2(\lambda) \tag{4·1}$$

as a measure of uncertainty depending on the form of the estimate $g_G(\lambda)$, and on the choice of r in $U_1^{(r)}$. We compare $g_D(\lambda)$ and $g_B(\lambda)$ by this criterion. The 'ideal' estimate (for $r = 2$) with the weighting function given in (3·4) is modified to
$$w_\lambda(\omega) = g(\lambda)/[2hg(\omega)] \tag{4·2}$$

if restricted to the interval $(\lambda - h, \lambda + h)$. For small h, $g(\omega) \sim g(\lambda)$, so that $g_D(\lambda)$ approximates to the 'ideal' estimate *if the latter is so restricted*. As, for $g_D(\lambda)$ with small h,
$$U_1^{(r)} \sim h(r+1)^{-1/r}, \quad U_2 \sim \tfrac{1}{2} g^2(\lambda)/h, \tag{4·3}$$

we have
$$W \sim \tfrac{1}{2}(r+1)^{-1/r}. \tag{4·4}$$

Table 1

$\lambda = \frac{1}{8}\pi$		$\lambda = \frac{1}{2}\pi$	
h: $\frac{1}{30}\pi$	0·290	h: $\frac{1}{10}\pi$	0·293
$\frac{1}{15}\pi$	0·292	$\frac{1}{6}\pi$	0·306
$\frac{1}{10}\pi$	0·301	$\frac{3}{10}\pi$	0·314
$\frac{2}{15}\pi$	0·311	$\frac{2}{3}\pi$	0·321
$\frac{1}{6}\pi$	0·322	$\frac{1}{2}\pi$	0·322

A numerical calculation (Table 1) was made of W for the restricted 'ideal' estimate ($r = 2$) for varying h in the particular case of the spectrum of M. G. Kendall's artificial series I (Kendall, 1946). As, for small h, $W \to 0{\cdot}289$ (for all λ *within* the total range), the above results show that W is smallest, in this particular case at least, as $h \to 0$, for the above class of estimate. More generally, for small h, it is readily shown that
$$W = \frac{1}{2\sqrt{3}} \left(1 + \frac{2ch^2}{15} \right), \tag{4·5}$$

where
$$c = \left(\frac{1}{g(\lambda)} \frac{dg(\lambda)}{d\lambda} \right)^2 - \frac{1}{2g(\lambda)} \frac{d^2 g(\lambda)}{d\lambda^2}, \tag{4·6}$$

so that whether W increases or decreases as h increases from zero depends on the sign of c; it will, however, increase in the neighbourhood of a maximum of $g(\lambda)$.

For $g_B(\lambda)$ it can be shown (e.g. by the use of (2·13)) that

$$\left.\begin{aligned} [U_1^{(2)}]^2 &\sim \frac{4}{n'}\left[\left(\frac{\pi-\lambda}{\pi}\right)\log\cos\tfrac{1}{2}\lambda+\frac{\lambda}{\pi}\log\sin\tfrac{1}{2}\lambda+\frac{1}{\pi}\log 2\right], \\ U_1^{(1)} &\sim \frac{2}{\pi n'}\log n', \\ [U_1^{(r)}]^r &\sim 1/[(n')^r\cos(\tfrac{1}{2}\pi r)\,\Gamma(2-r)] \quad (0<r<1). \end{aligned}\right\} \tag{4·7}$$

Making use of the result (2·17) (for $\lambda \neq 0$), we arrive at the comparison given in Table 2 of $g_D(\lambda)$ and $g_B(\lambda)$ in terms of W. Asymptotically, $g_D(\lambda)$ is thus superior to $g_B(\lambda)$ if the criterion W is taken for $r = 2$ (or 1), but the generalization of the resolvability measure U_1 emphasized its arbitrary character, and the 'efficiencies' of the two estimates cross over as r decreases.

Table 2

r	$g_D(\lambda)$	$g_B(\lambda)$
2	0·289	$0{\cdot}125\sqrt{n'}$ $(\lambda=\frac{1}{2}\pi)$
1	0·250	$0{\cdot}0675\log n'$
$\frac{1}{2}$	0·222	0·270
$\frac{1}{3}$	0·211	0·222
0·1	0·193	0·177
0+	0·184	0·162

5. Investigation of the theoretical optimum weighting function

The replacement of U_2 by $U_1^{(r)}U_2$ as a criterion of efficiency naturally complicates the investigation of the 'ideal' weighting function, but while this will still of course depend on a knowledge of the theoretical spectrum, it seemed worth while investigating the form of the optimum function if this criterion were adopted. We examine the case $r = 2$. Let $w_\lambda(\omega) \equiv v^2(\omega)$, where $v(\omega)$ is real, be the optimum function, and consider the slightly modified function $v(\omega)+\eta V(\omega)$, where η is small and $V(\omega)$ arbitrary. Denote the change in U_1U_2 by $\delta(U_1U_2)$. Now from (3·8)

$$U_1 = [(k_2-2\lambda k_1+\lambda^2 k_0)/k_0]^{\frac{1}{2}}, \tag{5·1}$$

and, to the first order of η,

$$\delta(k_n) = 2\eta K_n,$$

where

$$K_n \equiv \int_0^\pi \omega^n v(\omega)\,V(\omega)\,d\omega. \tag{5·2}$$

Hence

$$\delta(U_1) = \eta\{(K_2-2\lambda K_1+\lambda^2 K_0)/(U_1 k_0) - K_0 U_1/k_0\}.$$

Also

$$\delta(U_2') = 4\eta\int_0^\pi g^2(\omega)\,v^3(\omega)\,V(\omega)\,d\omega = 4\eta L, \quad \text{say.}$$

Thus finally

$$\delta(U_1U_2) = \frac{\eta}{k_0}\left[\frac{(K_2-2\lambda K_1+\lambda^2 K_0)\,U_2}{U_1} - K_0U_1U_2 + 4k_0U_1L\right]. \tag{5·3}$$

As $v(\omega)$ is assumed to be the optimum function, the coefficient of η should vanish. This implies that the expression within the square brackets in (5·3) should vanish for any appropriate $V(\omega)$. The condition of asymptotic unbiasedness gives the condition

$$\int_0^\pi g(\omega)\,v(\omega)\,V(\omega)\,d\omega = 0, \tag{5·4}$$

which will be satisfied if we take in particular

$$V(\omega)\,d\omega \equiv d\left[\frac{\epsilon(\omega-\alpha)}{g(\alpha)\,v(\alpha)} - \frac{\epsilon(\omega-\beta)}{g(\beta)\,v(\beta)}\right],$$

where $$\epsilon(x) \equiv \begin{cases} 0 & (x<0), \\ 1 & (x \geqslant 0). \end{cases}$$

The expression in square brackets in (5·3) now becomes

$$\frac{U_2}{U_1}\left[\frac{(\alpha-\lambda)^2}{g(\alpha)} - \frac{(\beta-\lambda)^2}{g(\beta)}\right] - U_1 U_2\left[\frac{1}{g(\alpha)} - \frac{1}{g(\beta)}\right] + 4k_0 U_1[g(\alpha)\,v^2(\alpha) - g(\beta)\,v^2(\beta)],$$

and, as this should be zero for arbitrary α, β, we obtain

$$\frac{U_2}{U_1}\frac{(\alpha-\lambda)^2}{g(\alpha)} - \frac{U_1 U_2}{g(\alpha)} + 4k_0 U_1 g(\alpha)\,v^2(\alpha) = \text{constant},$$

or

$$v^2(\alpha) = \frac{A}{g(\alpha)} + \frac{B}{g^2(\alpha)}\left[1 - \frac{(\alpha-\lambda)^2}{U_1^2}\right], \tag{5·5}$$

which determines the optimum form (if this exists in the above sense). It depends of course, as anticipated, on the unknown theoretical spectral function $g(\omega)$. In (5·5),

$$B = \tfrac{1}{2}U_2/k_0, \tag{5·6}$$

and the two formulae for U_1 and U_2 give, with the unbiasedness condition, three further relations for A, B, U_1 and U_2. If these yield an appropriate solution, the minimum $U_1 U_2$ is then also calculable.

6. Equations for the optimum function in the case of the second-order autoregressive process

The determination of the coefficients in (5·5) is unfortunately extremely laborious, and sometimes impossible, as was experienced in the particular case of the second-order autoregressive process

$$X_{t+2} + aX_{t+1} + bX_t = Y_{t+2}. \tag{6·1}$$

For (6·1) we have

$$g(\lambda) = 2\pi\sigma_X^2 f(\lambda) = \delta/(\alpha + \beta\cos\lambda + \gamma\cos 2\lambda), \tag{6·2}$$

where $\alpha = 1+a^2+b^2$, $\beta = 2a(1+b)$, $\gamma = 2b$, $\delta = 2\sigma^2$ (see Bartlett, 1950). Define

$$a_n^{(m)} = \int_0^\pi \omega^m g^{-n}(\omega)\,d\omega, \quad b_n^{(m)} = \int_0^\pi (\omega-\lambda)^m g^{-n}(\omega)\,d\omega,$$

which are evaluable for the $g(\lambda)$ of (6·2). Then it will be found that (5·6) yields the equation

$$\pi A^2 + 2A(a_1^{(0)}B - b_1^{(2)}C) + a_2^{(0)}B^2 - 2b_2^{(2)}BC + b_2^{(4)}C^2 = 4Bk_0, \tag{6·3}$$

where $C = B/U_1^2$, and

$$Aa_1^{(0)} + Ba_2^{(0)} - Cb_2^{(2)} = k_0. \tag{6·4}$$

The relation for U_1 give

$$Ab_1^{(2)} + Bb_2^{(2)} - Cb_2^{(4)} = k_0 B/C, \tag{6·5}$$

and the unbiasedness condition gives

$$A\pi + Ba_1^{(0)} - Cb_1^{(2)} = g(\lambda). \tag{6·6}$$

Substitution of k_0 in (6·3) and (6·5) gives, with the aid of (6·6), two equations of second degree in A and B. A being found in terms of B from one of them, a biquadratic equation in

A results from the other. An examination of an actual numerical case of (6·2) revealed that no solution in the admissible ranges for the unknowns is always possible; in any case, if we recall that the solutions even when they exist are only available when $g(\lambda)$ is known, these equations seem unlikely to be used in practice.

7. Theoretical gain in efficiency with weighting function in restricted range

In view of the difficulty of handling the optimum weighting function (on the basis of the criterion U), it may be advisable to demonstrate that the optimum function on the basis of the criterion U_2 but restricted to the *small* interval $\lambda-h$, $\lambda+h$, viz.

$$w(\lambda) = g(\lambda)/[2hg(\omega)] \sim 1/2h, \tag{7·1}$$

may at least be improved on the basis of the criterion U. As, for small h, $g(\omega) \sim g(\lambda)$ to the first order, the weighting function in (5·5) may be written to the same degree of approximation

$$A' + B'(\omega-\lambda)^2,$$

and it is easily found from the further relations for A, B, U_1 and U_2 that

$$w_W(\omega) \sim \frac{3}{4h}\left(1-\frac{(\omega-\lambda)^2}{h^2}\right); \tag{7·2}$$

Moreover, for (7·2), $$W = 3\sqrt{5}/25 = 0{\cdot}268.$$

This is only a small improvement over the value 0·289 for W with the Daniell estimate $g_D(\lambda)$, but it will be noticed that this approximate weighting function is independent of the unknown spectral function $g(\lambda)$, and hence also usable in practice. As the approximate bias in the resulting estimate $g_W(\lambda)$ is reduced from $\frac{1}{6}h^2d^2g(\lambda)/d\lambda^2$ to $\frac{1}{10}h^2d^2g(\lambda)/d\lambda^2$, the use of $g_W(\lambda)$ in place of $g_D(\lambda)$ might sometimes be preferable when the periodogram $I(\omega)$ is available.

The merits of the other estimate considered, $g_B(\lambda)$, rest in its asymptotic unbiasedness and its convenience of calculation from the first $n'-1$ sample autocovariances (or autocorrelations). From Table 2 we saw that its efficiency depended on what measure of resolvability was used.

REFERENCES

Bartlett, M. S. (1946). On the theoretical specification and sampling properties of autocorrelated time series. *J. Roy. Statist. Soc. Suppl.* **8**, 27.

Bartlett, M. S. (1950). Periodogram analysis and continuous spectra. *Biometrika*, **37**, 1.

Bartlett, M. S. (1954). Problèmes de l'analyse spectrale des series temporelles stationnaires. *Publ. de Inst. Statist. (Univ. de Paris)* vol. III, fasc. 3, 119.

Grenander, U. (1951). On empirical spectral analysis of stochastic processes. *Ark. Mat.* **1**, 503.

Grenander, U. & Rosenblatt, M. (1952). On spectral analysis of stationary time-series. *Proc. Nat. Acad. Sci., Wash.*, **38**, 519.

Grenander, U. & Rosenblatt, M. (1954). Statistical spectral analysis of time-series arising from stationary stochastic processes. *Ann. Math. Statist.* **24**, 537.

Kendall, M. G. (1946). *Oscillatory Time Series.* Cambridge University Press.

Slutsky, E. (1927). The summation of random causes as the source of cyclic processes. Problems of economic conditions, ed. by the Conjuncture Institute, Moscow, **3**, no. 1. (Later reprinted in *Econometrica*, **5** (1937), 105.)

Tukey, J. W. & Hamming, R. W. (1949). Measuring noise color, I. *Bell Tel. Lab. Memo.* MM–49–110.

Wold, H. (1938). *A Study in the Analysis of Stationary Time Series*, 1st ed. Uppsala.

THE AUTOCORRELATION FUNCTION AND THE SPECTRAL DENSITY FUNCTION

By J. WISE

London School of Economics

Introduction and summary

The relationship existing between the autocorrelation function and the spectral density function of a stationary, purely non-deterministic stochastic process is well known. Discussions of the relationship have been given by Wiener (1930) and Khintchine (1934), in the case of the continuous process, and by Wold (1954, Chapter 2, § 17, pp. 66–75), and Doob (1953, Chapter 10, §§ 3 and 4, pp. 473–86) for the discrete process. The treatment in the present paper is confined to the discrete process.*

The relationship referred to also provides a fundamental connexion between the autocorrelation matrix of a process and the corresponding spectral density function of the process. Some rather ingenious manipulations on the matrix representation of the spectral function have been carried out by Whittle (1951, Chapter 4, § 2, pp. 35–6, equations (4·272) and (4·276)). However, Whittle's work contains some inaccuracies. More important still, 'exact' results exist for those cases in which Whittle gave 'asymptotic' solutions.

In the course of the derivations of these exact relationships, the author uses methods which appear to be new—though bearing some relationship to André's method of solving linear difference equations (see C. Jordan, 1947, pp. 587–99)—and are capable of wide application in the theory of discrete linear stochastic processes. The use of exact relationships does, moreover, lead to a considerable improvement in clarity and rigour.

In view of the fundamental exact relationships existing between the autocovariance matrix and the spectral density function, the former might appropriately—in the case of the stationary process—be termed the spectral density matrix of the process. Traditional terminology has, however, been adhered to in the remainder of the paper.

The main topics treated below may be classified under the following headings.

(*a*) *The circular process*

(i) The exact relationship existing between the autocovariance matrix and the spectral density function.

(ii) The latent roots of the autocovariance matrix.

(iii) The exact relationship between the canonical form of the autocovariance matrix and the spectral density function.

(*b*) *The non-circular process*

(i) The exact relationship corresponding to (*a*) (i) above.

(ii) The exact inversion of non-circular autocovariance matrices.

* This discrete process may, however, be a sample realization drawn from a continuous process.

(*a*) The circular process

Definitions

The vector of N random variables, given by $\mathbf{x}' = \{x_N, x_{N-1}, \ldots, x_1\}$, is assumed to have the following distributional properties:

$$E\mathbf{x} = \mathbf{0}, \tag{1}$$

$$Ex_s^2 = \sigma^2 \quad \{s = 1, 2, \ldots, N\}, \tag{2}$$

$$Ex_s x_{s+L} = \sigma^2 \rho_L \quad \{s = 1, 2, \ldots, N\}, \tag{3}$$

where $$\rho_{N+L} = \rho_{N-L} = \rho_L.$$

The sequence of values $\rho_1, \rho_2, \ldots, \rho_N$ is termed the autocorrelation function of the process.

$$E\mathbf{x}\mathbf{x}' = \mathbf{V} \quad \text{(the autocovariance matrix)}$$

$$= \sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 & \ldots & \rho_2 & \rho_1 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \ldots & \rho_3 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 & \ldots & \rho_4 & \rho_3 \\ \vdots & & & & & & \vdots \\ \rho_1 & \rho_2 & \rho_3 & \rho_4 & \ldots & \rho_1 & 1 \end{bmatrix}, \tag{4}$$

an $N \times N$ matrix.

Properties

The autocovariance matrix, $\mathbf{V}$, may, from (4), be expressed in the following form:

$$\mathbf{V} = \sigma^2\{\mathbf{I} + \rho_1(\mathbf{W} + \mathbf{W}^{-1}) + \rho_2(\mathbf{W}^2 + \mathbf{W}^{-2}) + \ldots + \rho_{\frac{1}{2}(N-1)}(\mathbf{W}^{\frac{1}{2}(N-1)} + \mathbf{W}^{-\frac{1}{2}(N-1)})\}, \tag{5}$$

where N is a positive odd integer, or, alternatively,

$$= \sigma^2\{\mathbf{I} + \rho_1(\mathbf{W} + \mathbf{W}^{-1}) + \rho_2(\mathbf{W}^2 + \mathbf{W}^{-2}) + \ldots + \tfrac{1}{2}\rho_{\frac{1}{2}N}(\mathbf{W}^{\frac{1}{2}N} + \mathbf{W}^{-\frac{1}{2}N})\},$$

where N is a positive even integer. For all values of N, $\mathbf{I}$ denotes the $N \times N$ identity matrix, and $\mathbf{W}$ denotes the $N \times N$ circulant definition of the auxiliary identity matrix, that is to say,

$$\mathbf{W} = 0 \begin{bmatrix} 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & . & . & \ldots & 1 \\ 1 & 0 & . & . & \ldots & 0 \end{bmatrix}. \tag{6}$$

Since the matrices $(\mathbf{W}^L + \mathbf{W}^{-L})$ for $(L = 1, 2, \ldots, N)$ are commutative, they may all be simultaneously reduced to canonical form by a single transformation.

Thus, if we consider the matrix $\mathbf{L}$, satisfying the property

$$\mathbf{L}'\mathbf{L} = \mathbf{I}, \tag{7}$$

$$\mathbf{L}'(\mathbf{W}^L + \mathbf{W}^{-L})\mathbf{L} = \operatorname{diag}\{\omega_1^L + \omega_1^{-L}, \ldots, \omega_N^L + \omega_N^{-L}\}, \tag{8}$$

where $\omega_1, \omega_2, \ldots, \omega_N$ are the N values of the roots of the equation

$$y^N - 1 = 0, \tag{9}$$

then $$\mathbf{L}'\mathbf{V}\mathbf{L} = \operatorname{diag}\{\nu_1, \nu_2, \nu_3, \ldots, \nu_N\}$$

provides the canonical form of the autocovariance matrix $\mathbf{V}$.

Furthermore, we have

$$\omega_s^L + \omega_s^{-L} = 2\cos\frac{2\pi sL}{N} \quad (s = 1, 2, \ldots, N;\ L = 1, 2, \ldots, N). \tag{10}$$

The latent roots of V are thus given by

$$\begin{aligned} \nu_s &= \sigma^2\{1 + \rho_1(\omega_s + \omega_s^{-1}) + \rho_2(\omega_s^2 + \omega_s^{-2}) + \ldots + \rho_{\frac{1}{2}(N-1)}(\omega_s^{\frac{1}{2}(N-1)} + \omega_s^{-\frac{1}{2}(N-1)})\} \quad (s = 1, 2, \ldots, N) \\ &= \sigma^2\left\{1 + 2\rho_1\cos\frac{2\pi s}{N} + 2\rho_2\cos\frac{4\pi s}{N} + \ldots + 2\rho_{\frac{1}{2}(N-1)}\cos\frac{(N-1)\pi s}{N}\right\} \quad (s = 1, 2, \ldots, N) \end{aligned}$$

for N odd, or by

$$\begin{aligned} \nu_s &= \sigma^2\{1 + \rho_1(\omega_s + \omega_s^{-1}) + \rho_2(\omega_s^2 + \omega_s^{-2}) + \ldots + \tfrac{1}{2}\rho_{\frac{1}{2}N}(\omega_s^{\frac{1}{2}N} + \omega_s^{-\frac{1}{2}N})\} \quad (s = 1, 2, \ldots, N) \\ &= \sigma^2\left\{1 + 2\rho_1\cos\frac{2\pi s}{N} + 2\rho_2\cos\frac{4\pi s}{N} + \ldots + \rho_{\frac{1}{2}N}\cos\pi s\right\} \quad (s = 1, 2, \ldots, N) \end{aligned} \tag{11}$$

for N even.

The relationships given by (5) and (11) are the correct restatement of similar forms given by Whittle (1951. Chapter 4. § 2, pp. 35–6), which are incorrect even as approximations in the non-circular case.

General rules for the multiplicities of the latent roots of V have been incorrectly stated by both Whittle (1951, Chapter 4, § 3, p. 37) and Watson (1951*, Chapter 2, § 5, p. 50).

Results correctly obtained by R. L. Anderson (1942, § 5*a*, pp. 7–8), on the multiplicities in the values of the terms in the sequence,

$$\cos\frac{2\pi L}{N},\quad \cos\frac{4\pi L}{N},\quad \cos\frac{6\pi L}{N},\quad \ldots,\quad \cos 2\pi L, \tag{12}$$

may, however, be used to derive valid statements relating to the multiplicities in the latent roots of V.

Let $N = nk$ and $L = lk$, where n, l and k are positive integers, n and l being prime to one another. Two cases may then be distinguished.

(i) For n odd, there are $\frac{1}{2}(n-1)$ terms in the series (12) occurring with $2k$-fold multiplicities, together with the value unity occurring with k-fold multiplicity.

(ii) For n even, there are $\frac{1}{2}(n-2)$ terms in the series (12) occurring with $2k$-fold multiplicities, together with each of the values ± 1 occurring with k-fold multiplicities. For the derivation of these results the reader is referred to R. L. Anderson's paper.

Although it is impossible to enumerate here all the possibilities in detail, a 'rule' which the author has found useful in applications of the theory to sampling problems will be given relating to the multiplicities in the latent roots of V. This 'rule' is adequate for most of the cases which occur in practice. The 'rule' is not, however, a theorem, since exceptions to it occur when certain relationships exist between the elements of V. Care must therefore be taken when it is applied. This 'rule' is as follows:

If $\rho_L \neq 0$ for at least one value of L satisfying the condition $N = nk$ and $L = lk$, n and l being prime to one another, then the multiplicities in the latent roots will be at most $2k$-fold.

Whittle's assertion that if N is odd there are $\frac{1}{2}(N-1)$ pairs of distinct roots, with a single root greater than or less than all of these others; and if N is even there are $\frac{1}{2}(N-2)$ pairs of

* Watson states here that 'real symmetric circulant matrices have at most one odd latent root, the rest being pairwise equal'. That is, of course, incorrect.

distinct roots with one root greater than and one root less than all of the other roots, is incorrect.

An example occurring in the literature on time-series analysis, demonstrating a practical instance in which the 'rule' stated above yields the correct result, whereas Whittle's assertion breaks down, is the lag – L definition of the Markoff process, given by

$$x_t - \rho x_{t-L} = \epsilon_t \quad (t = 1, 2, \ldots, N), \tag{13}$$

where $x_{N+S} = x_S$, N and L not being prime to one another.

The spectral function of a circular process

The spectral density function of the circular process possessing the distributional properties (1), (2) and (3) is of the form*

$$\nu(\theta) = \sigma^2 \left\{1 + 2\rho_1 \cos\theta + 2\rho_2 \cos 2\theta + \ldots + 2\rho_{\frac{1}{2}(N-1)} \cos\frac{(N-1)\theta}{2}\right\} \quad \left\{\theta = \frac{2\pi}{N}, \frac{4\pi}{N}, \ldots, 2\pi\right\},$$

where N is odd; or of the form

$$\nu(\theta) = \sigma^2 \{1 + 2\rho_1 \cos\theta + 2\rho_2 \cos 2\theta + \ldots + \rho_{\frac{1}{2}N} \cos \tfrac{1}{2}N\theta\} \quad \left\{\theta = \frac{2\pi}{N}, \frac{4\pi}{N}, \ldots, 2\pi\right\}, \tag{14}$$

where N is even.

The latent roots of $\mathbf{V}$, given by (11), are thus the values of the spectral density function, $\nu(\theta)$, taken at constant intervals of $2\pi/N$ in the values of θ.

These roots may thus be regarded as being generated by the spectral density function of the process.

The results obtained so far are, of course, exact for finite N. By considering the limiting values of the roots as N tends to infinity, asymptotic results may be obtained.

The autoregressive-moving average process

The discrete circular linear process of the autoregressive-moving average type may be written in the form

$$x_t + \alpha_1 x_{t-1} + \ldots + \alpha_k x_{t-k} = \epsilon_t + \beta_1 \epsilon_{t-1} + \ldots + \beta_h \epsilon_{t-h} \quad \{t = 1, 2, \ldots, N\}, \tag{15}$$

where $x_{N+S} = x_S$, $\epsilon_{N+S} = \epsilon_S$, and

$$\left.\begin{aligned} E\epsilon_s &= 0 \quad \{s = 1, 2, \ldots, N\}, \\ E\epsilon_s^2 &= \sigma^2 \quad \{s = 1, 2, \ldots, N\}, \\ E\epsilon_s \epsilon_t &= 0 \quad \{s \neq t\}. \end{aligned}\right\} \tag{16}$$

* Any discrete stationary purely non-deterministic process may be written in the form

$$x_t = \eta_t + \beta_1 \eta_{t-1} + \beta_2 \eta_{t-2} + \ldots,$$

where the summation extends to infinity.

The spectral density function of this process is defined as

$$\nu(\theta) = (1 + \beta_1 e^{i\theta} + \beta_2 e^{2i\theta} + \ldots)(1 + \beta_1 e^{-i\theta} + \beta_2 e^{-2i\theta} + \ldots) \operatorname{var} \eta$$

for both circular and non-circular definitions of the process, for samples consisting of any number of consecutive observations. The relation (14) is obtained by expanding $\nu(\theta)$, thus defined, as a Fourier series.

In the matrix notation the process becomes

$$(\mathbf{I}+\alpha_1\mathbf{W}+\ldots+\alpha_k\mathbf{W}^k)\,\mathbf{x}=(\mathbf{I}+\beta_1\mathbf{W}+\ldots+\beta_h\mathbf{W}^h)\,\boldsymbol{\epsilon}, \tag{17}$$

where $\mathbf{x}'=\{x_N,x_{N-1},x_{N-2},\ldots,x_2,x_1\}$

and $\boldsymbol{\epsilon}'=\{\epsilon_N,\epsilon_{N-1},\epsilon_{N-2},\ldots,\epsilon_2,\epsilon_1\}$.

Solving (17) for $\mathbf{x}$, we obtain

$$\mathbf{x}=(\mathbf{I}+\alpha_1\mathbf{W}+\ldots+\alpha_k\mathbf{W}^k)^{-1}(\mathbf{I}+\beta_1\mathbf{W}+\ldots+\beta_h\mathbf{W}^h)\,\boldsymbol{\epsilon}. \tag{18}$$

Thus, from (18), we obtain $$E\mathbf{x}=\mathbf{0}, \tag{19}$$

and

$$\begin{aligned}E\mathbf{x}\mathbf{x}'=\mathbf{V}\quad&\text{(the autocovariance matrix)}\\&=\sigma^2(\mathbf{I}+\alpha_1\mathbf{W}+\ldots+\alpha_k\mathbf{W}^k)^{-1}(\mathbf{I}+\beta_1\mathbf{W}+\ldots+\beta_h\mathbf{W}^h)(\mathbf{I}+\beta_1\mathbf{W}^{-1}+\ldots+\beta_h\mathbf{W}^{-h})\\&\qquad\times(\mathbf{I}+\alpha_1\mathbf{W}^{-1}+\ldots+\alpha_k\mathbf{W}^{-k})^{-1}.\end{aligned} \tag{20}$$

The latent roots of $\mathbf{V}$ are therefore given by

$$\begin{aligned}\nu_s&=\sigma^2\frac{(1+\beta_1\omega_s+\ldots+\beta_h\omega_s^h)(1+\beta_1\omega_s^{-1}+\ldots+\beta_h\omega_s^{-h})}{(1+\alpha_1\omega_s+\ldots+\alpha_k\omega_s^k)(1+\alpha_1\omega_s^{-1}+\ldots+\alpha_k\omega_s^{-k})}\quad(s=1,2,\ldots,N)\\&=\sigma^2\frac{(1+\beta_1e^{2i\pi s/N}+\ldots+\beta_he^{2i\pi hs/N})(1+\beta_1e^{-2i\pi s/N}+\ldots+\beta_he^{-2i\pi hs/N})}{(1+\alpha_1e^{2i\pi s/N}+\ldots+\alpha_ke^{2i\pi ks/N})(1+\alpha_1e^{-2i\pi s/N}+\ldots+\alpha_ke^{-2i\pi ks/N})}\quad(s=1,2,\ldots,N).\end{aligned} \tag{21}$$

The spectral density function of the circulant process (15) may be written as

$$\begin{aligned}\nu(\theta)&=\sigma^2\frac{(1+\beta_1e^{i\theta}+\ldots+\beta_he^{ih\theta})(1+\beta_1e^{-i\theta}+\ldots+\beta_he^{-ih\theta})}{(1+\alpha_1e^{i\theta}+\ldots+\alpha_ke^{ik\theta})(1+\alpha_1e^{-i\theta}+\ldots+\alpha_ke^{-ik\theta})}\quad(0\leqslant\theta\leqslant2\pi)\\&=g(e^{i\theta})\quad\text{(let us say)}\quad(0\leqslant\theta\leqslant2\pi).\end{aligned} \tag{22}$$

It follows from (20) and (22) that $$\mathbf{V}=g(\mathbf{W}), \tag{23}$$

and, in addition, $$\nu_s=g(\omega_s)=g(e^{2i\pi s/N})\quad(s=1,2,\ldots,N). \tag{24}$$

These results are exact, and are fundamental to the exact treatment of circular processes.

An asymptotic relationship between $\mathbf{V}$ (in the non-circular case, to be further discussed in § (*b*) of this paper) and $g(\mathbf{W})$ has been given by Whittle (1951). It is not at all clear in what sense this asymptotic relationship is valid, unless the exact form is used. Whittle (1952, Theorem 1, (6) and (7), pp. 49–50), and Wold (1953, Chapter 11, § 2, Theorem 1) use the following interpretation of the asymptotic relationship:

$$\lim_{N\to\infty}\frac{\mathbf{x}'\mathbf{V}\mathbf{x}}{\mathbf{x}'g(\mathbf{W})\mathbf{x}}=1 \tag{25}$$

for almost all realizations of the process. This result, though valid, is difficult to work with rigorously. Whittle does not adhere to it in any of his derivations, and for a rigorous treatment it is preferable to work throughout with exact relationships. The latter are no more cumbersome than the asymptotic forms, and their use may lead to the avoidance of substantial errors.

This is abundantly illustrated in the case of the first-order autoregressive process in the circular case. The largest latent root of the autocorrelation matrix for $\alpha_1 < 0$ is given by $\frac{1-(-\alpha_1)^N}{1+(-\alpha_1)^N}\frac{1-\alpha_1}{1+\alpha_1}$. The corresponding asymptotic form of this is $\frac{1-\alpha_1}{1+\alpha_1}$, which differs substantially from the exact value when N is small and $-\alpha_1$ is close to unity. In fact, as $-\alpha_1$ tends to unity, the exact form tends in value to N, while the asymptotic form tends to infinity.

A final point which should be made on the subject of the latent roots of circulant variance matrices is that in the derivation of the values of, and of the multiplicities in these roots, it is usually more convenient to consider the spectral density function in the form (22) than in the form (14).

(*b*) The non-circular process

Whittle (1951, Chapter 4, § 2, relations (4·253), (4·262) and (4·28), pp. 34–6) used the 'asymptotic' form of (23) to obtain a simple method for the 'approximate' inversion of **V** in the non-circular case. For many purposes, approximate inversion is not adequate, and is not in any case necessary. It will be shown below that a very simple method of exact inversion is available. Whittle (1951, Chapter 4, § 2, relations (4·242)–(4·252), pp. 33–4) does give a cumbersome method of exact inversion based on the auxiliary identity matrix, denoted by **U**, where

$$\mathbf{U} = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{26}$$

an $N \times N$ matrix. Unfortunately Whittle's method appears to give the required result only in the case of the first-order autoregressive process.

An exact method of inversion will be given below which utilizes a non-circular counterpart of (20), the relation between the autocovariance matrix and the spectral density function.

The non-circular discrete linear process, generated by the relationship

$$x_t + \alpha_1 x_{t-1} + \dots + \alpha_k x_{t-k} = \epsilon_t + \beta_1 \epsilon_{t-1} + \dots + \beta_h \epsilon_{t-h}, \tag{27}$$

will be assumed to be semi-infinite. This does not prevent us from deriving results from samples consisting of N successive terms of the process. Such results for finite samples follow immediately from the case of the semi-infinite process. This is due to the fact that the autocovariance function and the spectral density function are both completely independent of N, a property which does not hold in the circular case.

We will assume (27) to represent a purely non-deterministic stationary process. An important theorem obtained by Wold (1954, Chapter 2, § 20, pp. 84–9, especially Theorem 7, p. 89) states, furthermore, that every semi-infinite, stationary, purely non-deterministic process can be expressed in the form (27).

A condition which is both necessary and sufficient for the stationarity of the process (27) can be obtained from a result given by Doob (1953).

The spectral density function of the process (27) is given by

$$g(e^{i\theta}) = \sigma^2 \frac{(1+\beta_1 e^{i\theta}+\ldots+\beta_h e^{ih\theta})(1+\beta_1 e^{-i\theta}+\ldots+\beta_h e^{-ih\theta})}{(1+\alpha_1 e^{i\theta}+\ldots+\alpha_k e^{ik\theta})(1+\alpha_1 e^{-i\theta}+\ldots+\alpha_k e^{-ik\theta})} \quad (0 \leqslant \theta \leqslant 2\pi).$$

It may, in passing, be noted that this is the same as (22), given for the circular process.

According to Doob (1953, Chapter 10, pp. 452–506, especially § 10, pp. 501–6), the necessary and sufficient condition for a stochastic process with the spectral function (22) to be stationary in the 'wide sense' (Chapter 2, § 8, pp. 94–5) is that the equation

$$z^h + \beta_1 z^{h-1} + \ldots + \beta_h = 0$$

shall have no roots outside the unit circle, and that the roots of the equation

$$z^k + \alpha_1 z^{k-1} + \ldots + \alpha_k = 0$$

shall all lie inside the unit circle. The reader is referred to Doob's* penetrating exposition for a more detailed treatment of this topic.

The relationship (27) may be transcribed as

$$(\mathbf{I}+\alpha_1 \mathbf{U}+\ldots+\alpha_k \mathbf{U}^k)\,\mathbf{x} = (\mathbf{I}+\beta_1 \mathbf{U}+\ldots+\beta_h \mathbf{U}^h)\,\boldsymbol{\epsilon}, \tag{28}$$

where $\mathbf{x}$ and $\boldsymbol{\epsilon}$ are semi-infinite vectors, and $\mathbf{U}$ is the semi-infinite auxiliary identity matrix.

Solving (28) for $\mathbf{x}$, leads to

$$\mathbf{x} = (\mathbf{I}+\alpha_1 \mathbf{U}+\ldots+\alpha_k \mathbf{U}^k)^{-1}(\mathbf{I}+\alpha_1 \mathbf{U}+\ldots+\beta_h \mathbf{U}^h)\,\boldsymbol{\epsilon}. \tag{29}$$

From (29) we deduce $$E\mathbf{x} = \mathbf{0}, \tag{30}$$

and $$\mathbf{V} = E\mathbf{x}\mathbf{x}'$$

$$= \sigma^2(\mathbf{I}+\alpha_1 \mathbf{U}+\ldots+\alpha_k \mathbf{U}^k)^{-1}(\mathbf{I}+\beta_1 \mathbf{U}+\ldots+\beta_h \mathbf{U}^h)(\mathbf{I}+\beta_1 \mathbf{U}'+\ldots+\beta_h \mathbf{U}'^h) \times (\mathbf{I}+\alpha_1 \mathbf{U}'+\ldots+\alpha_k \mathbf{U}'^k)^{-1}, \tag{31}$$

which is thus the autocovariance matrix of the process (27). This is the exact, non-circular counterpart to the relationships (20) and (23), given above for the case of the circular process.

Relationship (31) bears a close affinity to the 'autocovariance generating function' given by

$$g(z) = \sigma^2 \frac{(1+\beta_1 z+\ldots+\beta_h z^h)(1+\beta_1 z^{-1}+\ldots+\beta_h z^{-h})}{(1+\alpha_1 z+\ldots+\alpha_k z^k)(1+\alpha_1 z^{-1}+\ldots+\alpha_k z^{-k})}. \tag{32}$$

However, (31) is much more useful. From it can be obtained directly (i) the autocovariance function of the process (27), this being the property (31) shares with (32); (ii) the inverse of $\mathbf{V}$ in its exact form; (iii) the integral and rational powers of $\mathbf{V}$ and $\mathbf{V}^{-1}$.

Thus we have, from the elementary rule on the inversion of a product of several matrices,

$$\sigma^2\mathbf{V}^{-1} = (\mathbf{I}+\alpha_1 \mathbf{U}'+\ldots+\alpha_k \mathbf{U}'^k)(\mathbf{I}+\beta_1 \mathbf{U}'+\ldots+\beta_h \mathbf{U}'^h)^{-1} \times (\mathbf{I}+\beta_1 \mathbf{U}+\ldots+\beta_h \mathbf{U}^h)^{-1}(\mathbf{I}+\alpha_1 \mathbf{U}+\ldots+\alpha_k \mathbf{U}^k), \tag{33}$$

which is no more difficult to evaluate exactly than the matrix product in (31), which is the autocovariance matrix itself.

It is apparent from (31) and (33) that although $\mathbf{V}$ is a Laurent matrix, $\mathbf{V}^{-1}$ is not, a property already noted by Whittle.

* See also Wise (1955, Chapter III) for a further analysis of this problem.

The autoregressive process

To illustrate the application of the above results to problems of importance which have not been solved in the literature even in special cases, the inverse of the autocovariance matrix of the autoregressive process generated by the relation

$$x_t + \alpha_1 x_{t-1} + \ldots + \alpha_k x_{t-k} = \epsilon_t \tag{34}$$

will now be derived for the semi-infinite process.

From this result the inverses of the autocovariance matrices of N successive observations from the process (34) are then given in full for $k = 1$, $k = 2$ and $k = 3$.

From (33) we deduce for the special case (34) the result

$$\sigma^2 \mathbf{V}^{-1} = (\mathbf{I} + \alpha_1 \mathbf{U}' + \ldots + \alpha_k \mathbf{U}'^k)(\mathbf{I} + \alpha_1 \mathbf{U} + \ldots + \alpha_k \mathbf{U}^k). \tag{35}$$

Taking $k = 1$, 2 and 3 respectively, we deduce at once from (35) the matrices

(i) $k = 1$:

$$\sigma^2 \mathbf{V}^{-1} = \begin{bmatrix} 1 & \alpha_1 & 0 & \ldots & 0 & 0 \\ \alpha_1 & 1+\alpha_1^2 & \alpha_1 & \ldots & 0 & 0 \\ 0 & \alpha_1 & 1+\alpha_1^2 & \ldots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \ldots & 1+\alpha_1^2 & \alpha_1 \\ 0 & 0 & 0 & \ldots & \alpha_1 & 1 \end{bmatrix} \tag{36}$$

an $N \times N$ matrix.

(ii) $k = 2$:

$$\sigma^2 \mathbf{V}^{-1} = \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & 0 & \ldots & 0 & 0 \\ \alpha_1 & 1+\alpha_1^2 & \alpha_1+\alpha_1\alpha_2 & \alpha_2 & \ldots & 0 & 0 \\ \alpha_2 & \alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & \alpha_1+\alpha_1\alpha_2 & \ldots & 0 & 0 \\ 0 & \alpha_2 & \alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & \ldots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1+\alpha_1^2 & \alpha_1 \\ 0 & 0 & 0 & 0 & \ldots & \alpha_1 & 1 \end{bmatrix} \tag{37}$$

an $N \times N$ matrix.

(iii) $k = 3$:

$$\sigma^2 \mathbf{V}^{-1} = \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & \alpha_3 & \ldots & 0 & 0 \\ \alpha_1 & 1+\alpha_1^2 & \alpha_1+\alpha_1\alpha_2 & \alpha_2+\alpha_1\alpha_3 & \ldots & 0 & 0 \\ \alpha_2 & \alpha_1+\alpha_1\alpha_2 & 1+\alpha_1^2+\alpha_2^2 & \alpha_1+\alpha_1\alpha_2+\alpha_2\alpha_3 & \ldots & 0 & 0 \\ \alpha_3 & \alpha_2+\alpha_1\alpha_3 & \alpha_1+\alpha_1\alpha_2+\alpha_2\alpha_3 & 1+\alpha_1^2+\alpha_2^2+\alpha_3^2 & \ldots & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1+\alpha_1^2 & \alpha_1 \\ 0 & 0 & 0 & 0 & \ldots & \alpha_1 & 1 \end{bmatrix}. \tag{38}$$

The result (36) confirms that obtained for $k = 1$ by Cochrane & Orcutt (1949, equation (6·14), p. 57). However, the results (37) and (38) for $k = 2$ and $k = 3$ respectively do not appear to have been given previously in the literature.* So far as the author is aware, (35) is new, and the inverses of the autocovariance matrices for $k = 4, 5, 6, \ldots$, etc., can be written down immediately from it, rendering 'approximate' methods of inversion unnecessary.

* Note, however, Champernowne (1948, p. 206, equation 3·5), which gives the inverse implicitly for the case of the autoregressive process.

REFERENCES

ANDERSON, R. L. (1942). Distribution of the serial correlation coefficient. *Ann. Math. Statist.* **13**, 1.

CHAMPERNOWNE, D. G. (1948). Sampling theory applied to autoregressive schemes. *J. R. Statist. Soc.* **10**, 204.

COCHRANE, D. & ORCUTT, G. H. (1949). Application of least squares regression to relationships containing autocorrelated error terms. *J. Amer. Statist. Ass.* **44**, 32.

DOOB, J. L. (1953). *Stochastic Processes.* New York: John Wiley and Sons.

JORDAN, C. (1947). *Calculus of Finite Differences.* New York: Chelsea Publishing Co.

KHINTCHINE, A. (1934). Korrelationstheorie der stationaren stochastischen Prozesse. *Math. Ann.* **109**, 604.

WATSON, G. S. (1951). Serial correlation in regression analysis. Unpublished Ph.D. thesis and Inst. of Stat. Mimeo. Series, North Carolina State College.

WHITTLE, P. (1951). *Hypothesis Testing in Time-Series Analysis.* Uppsala: Almquist and Wiksell.

WHITTLE, P. (1952). Some results in time-series analysis. *Skand. ActuarTidskr.* **35**, 200.

WIENER, N. (1930). Generalized harmonic analysis. *Acta-Math.* **55**, 117.

WISE, J. (1955). Ph.D. Thesis (unpublished), University of London.

WOLD, H. (1953). *Demand Analysis.* New York: John Wiley and Sons.

WOLD, H. (1954). *A Study in the Analysis of Stationary Time Series.* Uppsala: Almquist and Wiksell.

SAMPLING PROPERTIES OF LOCAL STATISTICS IN STATIONARY STOCHASTIC SERIES

By G. H. JOWETT

University of Sheffield

1. Introduction

The sampling moments of large-sample statistics calculated from stationary time series usually require for their evaluation the summation of long series of cumulants or products of cumulants of the series, or for normal series the summation of long series of autocorrelations or products of autocorrelations (cf. Bartlett, 1946). These series usually converge only as the cumulants in question tend to zero, which is often in practice so slow a process as to make their evaluation a matter of considerable difficulty. There is, however, a class of statistics which may be called *local* in that they are dependent on short-term comparisons of terms in the series, and for these the summations are of such a kind as to converge as certain finite differences of the cumulants tend to zero; in practice this often happens much more quickly than the tending to zero of the cumulants themselves, and the corresponding formulae are accordingly much easier to evaluate in practice.

A simple illustration of this difference is provided by a comparison between the simplest statistic which depends on local comparisons and the simplest statistic which does not. It is advantageous to give this illustration before attempting a general discussion, so as to give a simple outline of the argument and the ideas involved.

Suppose $x_1, x_2, \ldots, x_{2n-1}, x_{2n}$ to be a set of $2n$ evenly spaced observations from a stationary time series with mean μ, variance σ^2, and autocorrelation function ρ_s. The sampling variance of the statistic

$$U \equiv n^{-1}[(x_1 - x_2) + (x_3 - x_4) + \ldots + (x_{2n-1} - x_{2n})] \tag{1}$$

is determined as follows:

$$\text{var } U \equiv n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \text{cov}\,[(x_{2i-1} - x_{2i}), (x_{2j-1} - x_{2j})] \tag{2}$$

$$= n^{-2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma^2(-\rho_{2(i-j)-1} + 2\rho_{2(i-j)} - \rho_{2(i-j)+1}) \tag{3}$$

$$= n^{-1} \sum_{s=-(n-1)}^{+(n-1)} \left(1 - \frac{|s|}{n}\right)(-\Delta''\sigma^2\rho_{2s}), \tag{4}$$

where Δ'' is the second central difference operator in Comrie's notation, and refers to the suffix of ρ. The need for this statistic and its sampling variance might arise, for example, in the estimation of the treatment effect in a systematic experiment having the pattern

$$T\ C\ T\ C\ \ldots\ T\ C \quad (T \equiv \text{treatment};\ C \equiv \text{control}), \tag{5}$$

or, alternatively, in the statistical control of the difference between the means of two interpenetrating systematic samples (Jowett, 1955*b*).

On the other hand, the sampling variance of the simple arithmetic mean

$$T \equiv (n^{-1})(x_1 + x_3 + \ldots + x_{2n-1}) \tag{6}$$

is easily shown to be given by the formula

$$\operatorname{var} T = \frac{1}{n} \sum_{s=-(n-1)}^{+(n-1)} \left(1 - \frac{|s|}{n}\right) \sigma^2 \rho_{2s}. \tag{7}$$

In many applications, as $|s|$ increases the autocorrelation function tends to become 'locally linear', i.e. approximately linear over intervals of width equal, say, to twice the interval between successive terms of the series; as $|s|$ increases further it gradually tends to zero. Thus $\Delta'' \rho_{2s}$ may attain a sufficient degree of smallness to be neglected in computing the approximate value of the sum in (4) at a much earlier stage in the increase of $|s|$ than that for which ρ_{2s} can be neglected in computing the approximate value of the sum in (7).

Now experience and *a priori* considerations (cf. Jowett, 1953) suggest that for many series observed in practice the graph of the serial variation function

$$\delta_s = E[\tfrac{1}{2}(x_i - x_{i+s})^2] \tag{8}$$

is zero for $s = 0$, rises steeply as s increases with gradually decreasing gradient and curvature, and ultimately flattens out when s becomes large, rather like the graph of the function

$$y = A(1 - e^{-ks})$$

for positive values of the constants A and k.

Since
$$\delta_s = \sigma^2(1 - \rho_s), \tag{9}$$

it follows that
$$-\Delta'' \sigma^2 \rho_s = \Delta'' \delta_s, \tag{10}$$

and hence that the graph of $\sigma^2 \rho_s$ will have the same shape but inverted. The serial variation function has the advantage that for short lags it is effectively independent of long-term variation in the series and requires no reference to the series mean, qualities which make it eminently suitable for use in the study of local variation. If for $s \geqslant s_0(\epsilon)$ the serial variation function has become approximately linear to such an extent that $\Delta'' \delta_{2s}$ is less than some predetermined small quantity ϵ, we deduce from (4) that

$$\operatorname{var} U = \frac{1}{n} \sum_{s=-(s_0-1)}^{+(s_0-1)} \left(1 - \frac{|s|}{n}\right) \Delta'' \delta_{2s} + R, \tag{11}$$

where
$$|R| < \frac{2\epsilon}{n} \sum_{s=s_0}^{(n-1)} 1 - \frac{|s|}{n} < \epsilon. \tag{12}$$

Hence the error in the estimate of var U committed by truncating the summation at $\pm(s_0 - 1)$ is at worst of the same order of magnitude as the departure from local linearity of δ_s over the rest of the range, which is usually the same as that of the first term omitted from the summation, since $|\Delta'' \delta_{2s}|$ will usually tend to decrease monotonically for $s > s_0$.

Since we are usually obliged to infer the stochastic properties from one or at most a very few realizations of the series itself, it will often be true that var U is much easier to estimate than var T; the need to estimate the form of ρ_s for large lags, often a matter of considerable difficulty, is not present, and the formula (4) may be computed approximately (provided, of course, that $\epsilon = o(n^{-1})$ and $s_0 \ll n$) using serial variation statistics with relatively short lags, which are usually readily available.

It will be observed that U is a function of the set of differences $x_1 - x_2$, $x_3 - x_4$, ..., which might be described as *local comparisons*, since they are constructed from adjacent terms of

the series. It is this which gives rise to the differencing in (4), which makes the formula independent of σ^2, and hence expressible entirely in terms of δ_s, and which enables var U to be estimated in many circumstances where it would be impossible to estimate var T. A corresponding property has been observed in other statistics, such as the mean of a systematics ample (cf. Jowett, 1952) or a trend-reduced regression coefficient (Jowett, 1955*a*) which are functions of local comparisons alone, and suggests that the sampling properties of statistics which are built up entirely from local comparisons are essentially determined by the local variational properties (e.g. δ_s for small s) of the series themselves, and are effectively independent of the long-term variational properties of the series which are so difficult to measure in practice. The rest of this paper will be concerned with establishing this principle, and with illustrating it by some specific formulae.

2. Sampling moments of local statistics in stationary normal series

We shall assume a reference space S with position vector $\mathbf{t}$, and shall use the symbol $\boldsymbol{\tau}$ to denote a displacement vector, e.g. $\mathbf{t}_1 - \mathbf{t}_2$. This is a generalization of the usual time axis in one dimension. A set of stochastic variables $x_1(\mathbf{t}), x_2(\mathbf{t}), \ldots$ (which may be regarded as components of a vector random function) with respective means $\mu_1, \mu_2, \ldots$ and variances $\sigma_1^2, \sigma_2^2, \ldots$, is defined at the points of S; in practice the definition is often at a lattice of points only, but for generality we shall assume definition at all points of S. These variables will be taken to form stationary series, i.e. their probability parameters to be invariant under translation in S.

We shall be concerned with *seminvariant local linear functions* (s.l.l.f.'s) of these variables defined as follows:

$$L_{xi} \equiv \int_s x_i(\mathbf{t})\, dL_x(\mathbf{t}), \quad \int_s dL_x(\mathbf{t}) = 0. \tag{13}$$

The integration is to be understood in the Stieltjes sense, so that

$$L_{xi} \equiv \sum_r l^*_{xr} x_i(\mathbf{t}_r) + \int_s x_i(\mathbf{t})\, l^{**}_x(\mathbf{t})\, d\mathbf{t}, \tag{14}$$

where l^*_{xr} is zero except at a finite number of points of S, and $l^{**}(\mathbf{t})$ is integrable in the ordinary Riemann sense. These linear functions of $x(\mathbf{t})$ we called *seminvariant* because they are unchanged by the addition of any constant to $x(\mathbf{t})$; to justify the adjective *local*, l^*_{xr} and $l^{**}_x(\mathbf{t})$ will be taken as zero outside a minimal sphere in S of diameter a_x bounded above by some fixed value a. These s.l.l.f.'s are generalizations of the simple local comparisons implied in § 1.

It will be assumed that the series are multivariate normal. This assumption may, however, be dispensed with when we are concerned only with second moments of linear statistics, since the question of normality does not then arise. The assumption of normality implies that

$$E x_1(\mathbf{t}_1) x_2(\mathbf{t}_2) \ldots x_s(\mathbf{t}_s) = \begin{cases} \mu_1\mu_2 \ldots \mu_s + \text{other terms each having at least one of } \mu_1\mu_2 \ldots \mu_s \\ \quad \text{as a factor} \quad (s \text{ odd}), \\ \mu_1\mu_2 \ldots \mu_3 + \sum\limits_{\alpha\beta\ldots\nu} \operatorname{cov}_{x_\alpha x_\beta}(\mathbf{t}_\alpha - \mathbf{t}_\beta) \operatorname{cov}_{x_\gamma x_\delta}(\mathbf{t}_\gamma - \mathbf{t}_\delta) \ldots \operatorname{cov}_{x_\mu x_\nu}(\mathbf{t}_\mu - \mathbf{t}_\nu) \\ \qquad + \text{other terms each having at least one of } \mu_1\mu_2 \ldots \mu_s \\ \qquad\quad \text{as a factor} \quad (s \text{ even}), \end{cases} \tag{15}$$

where $\alpha\beta \ldots \mu\nu$ is any permutation of $1 \ldots s$ such that

$$\alpha < \beta,\ \gamma < \delta,\ \ldots,\ \mu < \nu \quad \text{and} \quad \alpha < \gamma \ldots < \mu. \tag{16}$$

If we prefer to work with lag variation parameters defined by

$$\delta_{x_i x_j}(\mathbf{t}-\mathbf{t}') = E\,\tfrac{1}{2}(x_i(\mathbf{t})-x_j(\mathbf{t}'))^2, \tag{17}$$

we have

$$\delta_{x_i x_j}(\mathbf{t}-\mathbf{t}') = \tfrac{1}{2}(\mu_i-\mu_j)^2+\tfrac{1}{2}(\sigma_i^2+\sigma_j^2)-\operatorname{cov}_{x_i x_j}(\mathbf{t}-\mathbf{t}') \tag{18}$$

as the relation which permits us to substitute δ for cov. In analysis concerned with s.l.l.f.'s, the seminvariance usually results in the elimination of the terms involving means and variances. For example,

$$\begin{aligned}\operatorname{cov}(L_{\alpha i}, L_{\beta j}) &= E(L_{\alpha i}L_{\beta j}) \quad (\text{since } E(L_{\alpha i})=0)\\ &= \iint_S x_i(\mathbf{t})\,x_j(\mathbf{t}')\,dL_\alpha(\mathbf{t})\,dL_\beta(\mathbf{t}')\\ &= \iint_S \operatorname{cov}_{x_i x_j}(\mathbf{t}-\mathbf{t}')\,dL_\alpha(\mathbf{t})\,dL_\beta(\mathbf{t}')\\ &= \iint_S -\delta_{x_i x_j}(\mathbf{t}-\mathbf{t}')\,dL_\alpha(\mathbf{t})\,dL_\beta(\mathbf{t}').\end{aligned} \tag{19}$$

The double integration is interpretable in the obvious way in terms of summations and Riemann integrations.

In addition to the assumption of normality, we shall make a further important assumption about the nature of the lag variation functions (l.v.f.'s), an assumption which is very often reasonable in practice (cf. Jowett, 1953). We shall assume that the l.v.f.'s tend to become increasingly linear as $|\tau|$ increases; or, more precisely, that for arbitrary (small) $\epsilon>0$, and for arbitrary c, there is a value $h_0(\epsilon, c)$ such that for $h>h_0$ and $h<|\boldsymbol{\tau}|<h+c$, we can find a constant ψ_{ij} and a constant vector $\boldsymbol{\chi}_{ij}$ such that

$$\delta_{x_i x_j}(\boldsymbol{\tau}) = \psi_{ij}+\boldsymbol{\chi}_{ij}.\boldsymbol{\tau}+\epsilon_{ij}(\boldsymbol{\tau}), \tag{20}$$

where

$$|\epsilon_{ij}(\boldsymbol{\tau})|<\epsilon. \tag{21}$$

Two s.l.l.f.'s will be described as *far apart* if their associated minimal spheres touch the outside of a sphere of diameter greater than some predetermined distance h_0. If we take h_0 as dependent on ϵ as described above, and take c as twice the maximum local diameter a, it follows that if L_α, L_β are far apart,

$$\begin{aligned}\operatorname{cov}(L_{\alpha i}, L_{\beta j}) &= \iint -\delta_{x_i x_j}(\mathbf{t}-\mathbf{t}')\,dL_\alpha(\mathbf{t})\,dL_\beta(\mathbf{t}')\\ &= -\iint[\psi_{ij}+\boldsymbol{\chi}_{ij}.(\mathbf{t}-\mathbf{t}')+\epsilon_{ij}(\mathbf{t}-\mathbf{t}')]\,dL_\alpha(\mathbf{t})\,dL_\beta(\mathbf{t}')\\ &= -\iint\epsilon_{ij}(\mathbf{t}-\mathbf{t}')\,dL_\alpha(\mathbf{t})\,dL_\beta(\mathbf{t}'),\end{aligned} \tag{22}$$

and hence that

$$\begin{aligned}|\operatorname{cov}(L_{\alpha i}, L_{\beta j})| &\leqslant \epsilon\int_S |dL_\alpha(\mathbf{t})|\int_S |dL_\beta(\mathbf{t})|\\ &= O(\epsilon).\end{aligned} \tag{23}$$

On the other hand, if L_α and L_β are not far apart, $\operatorname{cov}(L_{\alpha i}, L_{\beta j})$ is a function of $\delta_{x_i x_j}(\boldsymbol{\tau})$ over the range

$$0\leqslant|\boldsymbol{\tau}|<h_0+2a, \tag{24}$$

i.e. in what may be called a (h_0+2a) neighbourhood of $\tau = 0$. If h_0 and a are sufficiently small (in practice, h_0 is often comparable with a in magnitude), the behaviour of $\delta_{x_i x_j}(\tau)$ in such a neighbourhood may be interpreted as a local variational property of the series.

This property of being either dependent on the l.v.f. in a neighbourhood of $\tau = 0$ or of magnitude $O(\epsilon)$ is also true of the cross-moment of several s.l.l.f.'s. For the cross-moment of s of these functions, which have zero expectations, we have

$$E(L_{1i_1}, L_{2i_2} \dots L_{si_s}) = \iint \dots \int E[x_{i_1}(\mathbf{t})\, x_{i_2}(\mathbf{t}') \dots x_{i_s}(\mathbf{t}^{(s-1)})]\, dL_1(\mathbf{t})\, dL_2(\mathbf{t}') \dots dL_s(\mathbf{t}^{(s-1)})$$

$$= \begin{cases} 0 & \text{if } s \text{ is odd,} \\ \iint \dots \int (-1)^{\frac{1}{2}s} \sum_{\alpha\beta\dots\mu\nu} [\delta_{(\alpha\beta)}(\mathbf{t}-\mathbf{t}')\, \delta_{(\gamma\delta)}(\mathbf{t}''-\mathbf{t}''') \dots \delta_{(\mu\nu)}(\mathbf{t}^{(s-2)}-\mathbf{t}^{(s-1)})] \\ \quad \times dL_\alpha(\mathbf{t}) \dots dL_\nu(\mathbf{t}^{(s-1)}) \text{ if } s \text{ is even,} \end{cases} \tag{25}$$

(where the suffix $(\alpha\beta)$ means $x_{i_\alpha} x_{i_\beta}$), since all the terms in the expectation of the product in the square bracket which have one or more of $\mu_1 \dots \mu_s$ as factor are annihilated in the integration by the seminvariant property of the corresponding L's.

The summation in (25) is taken over the set of permutations of $1, 2, \dots, s$ defined in (16). Hence if s is even,

$$E(L_{1i_1} \dots L_{si_s}) = \sum_{\alpha\beta\dots\nu} \operatorname{cov}(L_{\alpha i_\alpha}, L_{\beta i_\beta}) \operatorname{cov}(L_{\gamma i_\gamma}, L_{\delta i_\delta}) \dots \operatorname{cov}(L_{\mu i_\mu}, L_{\nu i_\nu}). \tag{26}$$

Since the right-hand side of (26) is a function of covariances, it follows that $E(L_{1i_1} L_{2i_2} \dots L_{si_s})$ is either dependent on the l.v.f. in a neighbourhood of $\tau = 0$ or of magnitude at most of order ϵ.

If the series $x_{i_1}(\mathbf{t}), \dots$ are made to coincide in sets, and the s.l.l.f.'s $L_{1i_1}, L_{2i_2}, \dots$ to coincide in sets included in them, the result just established takes the following form:

Theorem. The expectation of any product of powers of s.l.l.f.'s of any set of concomitant stationary stochastic variables for which the s.v.f.'s and the l.v.f.'s have a tendency towards linearity with increasing $|\tau|$ (as defined by (20) and (21) and the preceding paragraph) is a sum of terms which are either of magnitude at most of order ϵ or involve only values of the s.v.f.'s and l.v.f.'s in a (h_0+2a) neighbourhood of $\tau = 0$.

It follows that the moments and cumulants of any statistic which is a product of powers of s.l.l.f.'s have the same property, since these are expressible as linear functions of products of powers; and hence that the moments and cumulants of any statistic which can be expressed as a linear function of products of powers has the same property. The theorem is thus seen to be very general in character, for most if not all statistics which measure local properties and are of interest in practice may be expressed either exactly or approximately in this form. The theorem implies, broadly speaking, that the variational properties of local statistics depend essentially on local variational properties of the parent series.

3. Asymptotic moments of large-sample local statistics

Suppose that the statistics $L_{1i_1} \dots L_{si_s}$ (s even and divisible by an integer p) fall into r sets of p, namely, $\mathscr{M}_1 \dots \mathscr{M}_r$. Suppose, moreover, that the product of the statistics falling into the set $\mathscr{M}_i$ is denoted by M_i. Thus

$$M_1 = L_{1i_1} L_{2i_2} \dots L_{pi_p}, \quad M_2 = L_{p+1, i_{p+1}} L_{p+2, i_{p+2}} \dots L_{2p, i_{2p}}, \quad \dots, \quad M_r = L_{\overline{r-1}p+1, i_{\overline{r-1}p+1}} \dots L_{s, i_s}. \tag{27}$$

It is easily shown that

$$E(M_1 - E(M_1))(M_2 - E(M_2)) \dots (M_r - E(M_r)) = \sum_{\alpha\beta\dots\mu\nu} \operatorname{cov}(L_{\alpha i_\alpha} L_{\beta i_\beta}) \operatorname{cov}(L_{\gamma i_\gamma} L_{\delta i_\delta}) \dots \operatorname{cov}(L_{\mu i_\mu} L_{\nu i_\nu}), \quad (28)$$

where $\alpha\beta \dots \mu\nu$ is a permutation of the integer $l \dots s$, subject to the conditions (16) and also to the condition that no pair of suffixes associated with a covariance included in the formula may come from the same set $\mathscr{M}$.

Two sets will be said to be *far apart* if the members of one set are far apart from the members of the other. If any set is far apart from all the others, the product-moment (28) will be $O(\epsilon^p)$, since every member of the set has to be associated in a covariance with some member of another set. For some of the terms in (28) not to be $O(\epsilon^p)$, for each set there must be at least one other set from which it is not far apart; if this is so, terms which are not $O(\epsilon^p)$ arise, but only when members from such neighbouring sets occur together in the covariance brackets.

Many statistics which are of practical interest are means, or functions of means, of powers or products of powers of s.l.l.f.'s of the stochastic variables over a sample region of the space S or at a set of sample points, usually evenly spaced. Suppose that we are given sets $\mathscr{M}_1, \mathscr{M}_2, \dots, \mathscr{M}_n$, where n is large (for example, of the same order of magnitude as the number of sample points), and where the sets are *evenly spread* over the sample region in such a way that every subregion of volume $O(V/n)$ (i.e. $O(1)$) contains sets to the number $O(1)$. Then if

$$U = (n^{-1})(M_1 + M_2 + \dots + M_n), \quad (29)$$

we have

$$E(U - E(U))^r = (n^{-r}) E \sum_{\alpha_1 \dots \alpha_r} (M_{\alpha_1} - E(M_{\alpha_1})) \dots (M_{\alpha_r} - E(M_{\alpha_r})), \quad (30)$$

where $\alpha_1 \dots \alpha_r$ is a permutation of r of the integers $1 \dots n$.

If we assume $h \ll n$, the number of terms in (30) which are not $O(\epsilon^p)$, i.e. which can be formed from products of covariances of s.l.l.f.'s which are not far apart, is $O(n^{\frac{1}{2}r})$, since the number of ways of choosing pairs of sets from $\mathscr{M}_1 \mathscr{M}_2 \dots \mathscr{M}_n$ which are not far apart is of this order. Hence

$$E(U - E(U))^r = O(n^{-\frac{1}{2}r}) + O(\epsilon^p), \quad (31)$$

the second term on the right-hand side of (31) being justified by the fact that the summation in (30) has only n^r terms altogether, and is divided by n^r.

If $\epsilon^p = o(n^{-\frac{1}{2}r})$, the moment will be dominated by those terms which involve only values of the l.v.f. in a $(h + 2a)$ neighbourhood of $\tau = 0$, and will be of magnitude $O(n^{-\frac{1}{2}r})$.

This argument may be generalized in a fairly obvious way to show that cross-moments of order r about the mean for statistics having the same form as U also have this property and this order of magnitude. Thus, since many statistics of useful potential application have sampling errors which may be expressed, to a large-sample approximation at least, as linear functions of products of sampling errors of statistics having this form, it may be shown that the lower moments of these also depend essentially on values of the l.v.f. in a restricted neighbourhood of $\tau = 0$.

4. EXAMPLES

Example 1. *Variance of a trend-reduced covariance*

Suppose we have a sample stretch $x_1 y_1, x_2 y_2, \dots, x_n y_n$ from the variation of two concomitant series, x, y. The covariance of the difference between successive terms, which is the simplest case of a trend-reduced covariance (cf. Jowett, 1955*a*), is defined by the equation

$$U = (n-1)^{-1}[(x_1 - x_2)(y_1 - y_2) + \dots + (x_{n-1} - x_n)(y_{n-1} - y_n)]. \quad (32)$$

Then

$$\begin{aligned}\operatorname{var} U &= (n-1)^{-2}\sum_{i,j=1}^{n-1}\operatorname{cov}\left[(x_i-x_{i+1})(y_i-y_{i+1}),(x_j-x_{j+1})(y_j-y_{j+1})\right]\\ &= (n-1)^{-2}\sum_{i,j=1}^{n-1}\operatorname{cov}\left[(x_i-x_{i+1}),(x_j-x_{j+1})\right]\operatorname{cov}\left[(y_i-y_{i+1}),(y_j-y_{j+1})\right]\\ &\qquad +\operatorname{cov}\left[(x_i-x_{i+1}),(y_j-y_{j+1})\right]\operatorname{cov}\left[(y_i-y_{i+1}),(x_j-x_{j+1})\right]\\ &= (n-1)^{-2}\sum_{i,j=1}^{n-1}\Delta''\delta_{xx}(i-j)\,\Delta''\delta_{yy}(i-j)+[\Delta''_{xy}(i-j)]^2\\ &= (n-1)^{-1}\sum_{s=-(n-2)}^{+(n-2)}\left(1-\frac{|s|}{n-1}\right)\{\Delta''\delta_{xx}(s)\,\Delta''\delta_{yy}(s)+[\Delta''_{xy}(s)]^2\}. \qquad (33)\end{aligned}$$

For independent series, $\delta_{xy}(s)$ is constant for all s, so that

$$\Delta''\delta_{xy}(s)=0. \qquad (34)$$

The dominant terms in the summation are those for small $|s|$; as $|s|$ increases, and the δ's straighten out, the remaining terms will rapidly tend to zero.

Example 2. The covariance of two serial variation statistics

The serial variation statistics of lags s_1 and s_2 $(s_1 \leqslant s_2)$ calculated from a stretch of series $x_1 \ldots x_n$ are defined as follows:

$$\begin{aligned}d(s_1) &= (n-s_1)^{-1}\sum_{i=1}^{n-s_1-1}\tfrac{1}{2}(x_i-x_{i+s_1})^2,\\ d(s_2) &= (n-s_2)^{-1}\sum_{i=1}^{n-s_2-1}\tfrac{1}{2}(x_i-x_{i+s_2})^2. \qquad (35)\end{aligned}$$

Then

$$\begin{aligned}&\operatorname{cov}\left(d(s_1),d(s_2)\right)\\ &\quad=(n-s_1)^{-1}(n-s_2)^{-1}\sum_{i=1}^{n-s_1-1}\sum_{j=1}^{n-s_2-1}\operatorname{cov}\left(\tfrac{1}{2}(x_i-x_{i+s_1})^2,\tfrac{1}{2}(x_j-x_{j+s_2})^2\right)\\ &\quad=(n-s_1)^{-1}(n-s_2)^{-1}\sum_{i,j}2[\operatorname{cov}\left(\tfrac{1}{2}(x_i-x_{i+s_1}).\tfrac{1}{2}(x_j-x_{j+s_2})\right)]^2\\ &\quad=(n-s_1)^{-1}(n-s_2)^{-1}\sum_{\alpha=-(n-s_2-1)}^{n-s_1-1}(n-c_\alpha).\tfrac{1}{2}[-\delta(\alpha)-\delta(\alpha+s_2-s_1)+\delta(\alpha-s_1)+\delta(\alpha+s_2)]^2\\ &\quad=n^{-1}\sum_{\alpha=-(n-s_2-1)}^{n-s_1-1}\frac{n(n-c_\alpha)}{(n-s_1)(n-s_2)}.\tfrac{1}{2}[-\delta(\alpha)-\delta(\alpha+s_2-s_1)+\delta(\alpha-s_1)+\delta(\alpha+s_2)]^2, \qquad (36)\end{aligned}$$

where

$$c_\alpha=\begin{cases}s_1+|\alpha|, & s_2-s_1\leqslant\alpha\leqslant n-s_1-1,\\ s_2, & 0\leqslant\alpha\leqslant s_2-s_1,\\ s_2+|\alpha|, & -(n-s_2-1)\leqslant\alpha\leqslant 0.\end{cases} \qquad (37)$$

The dominant terms in the summation in (36) are again those for small $|s|$. As $|s|$ increases, and $\delta(s)$ straightens out, the remaining terms, being proportional to squares of a kind of second difference, again tend rapidly to zero.

Example 3. Spatial systematic sampling

The systematic sampling will be of the variable $x(u,v)$ which is defined over the rectangular region $0\leqslant u\leqslant r_1$, $0\leqslant v\leqslant r_2$, samples being taken at the points $(i-\frac{1}{2}, j-\frac{1}{2})$ for integral i, j such that $i=1,\ldots,r_1$, $j=1,\ldots,r_2$.

Let

$$L_{ij}=x(i-\tfrac{1}{2},j-\tfrac{1}{2})-\int_{u=i-1}^{i}\int_{v=j-1}^{j}x(u,v)\,du\,dv. \qquad (38)$$

Then

$$\operatorname{cov}(L_{ij}, L_{i'j'}) = -\delta(i-i', j-j') - \int_{\theta=i-i'-1}^{i-i'+1}\int_{\phi=j-j'-1}^{j-j'+1} (1-|\theta-i+i'|)(1-|\phi-j+j'|)\,\delta(\theta,\phi)\,d\theta\,d\phi$$

$$+\int_{i'-1}^{i'}\int_{j'-1}^{j'} \delta(i-\tfrac{1}{2}-u, j-\tfrac{1}{2}-v)\,du\,dv + \int_{i-1}^{i}\int_{j-1}^{j} \delta(u-i'-\tfrac{1}{2}, v-j'-\tfrac{1}{2})\,du\,dv. \quad (39)$$

This covariance is itself a s.l.l.f. of $\delta(\theta,\phi)$ over the region

$$i-i'-1 \leqslant \theta \leqslant i-i'+1, \quad j-j'-1 \leqslant \phi \leqslant j-j'+1, \quad (40)$$

and will be negligible if the points (i,j), (i',j') are far enough apart.

The sampling variance of the mean U of a systematic sample of $n\,(=r_1r_2)$ about the mean of the rectangular region is therefore given by

$$\operatorname{var} U \equiv \operatorname{var}\left(n^{-1}\sum_{i=1}^{r_1}\sum_{j=1}^{r_2} L_{ij}\right) \quad (41)$$

$$= n^{-2}\sum_{ij,i'j'=1}^{n} \operatorname{cov}(L_{ij}, L_{i'j'}) \equiv (n^{-2})\sum_{ij,i'j'} c(i-i', j-j'), \quad \text{say.} \quad (42)$$

In evaluating the summation in (42), we take the terms in order of increasing distance between (i,j), (i',j'). Hence

$$\operatorname{var} U = \frac{1}{n}\left\{c(0,0) + \left[\frac{2n}{(r_1-1)\,r_2}\,c(1,0) + \frac{2n}{r_1(r_2-1)}\,c(0,1)\right] + \frac{4n}{(r_1-1)(r_2-1)}\,c(1,1)\ldots\right\}. \quad (43)$$

As the distance increases and $\delta(\theta,\phi)$ becomes locally linear, the terms will become negligible.

In Ex. 1, 2 and 3 we are dealing with second moments, corresponding to $r = 2$ in (30). Hence we verify that they are all of order $n^{-ir} \equiv n^{-1}$, provided that local linearity is attained rapidly enough. The next example is different, involving a greater value of r.

Example 4. *Fourth moment of the statistic* $U \equiv (n^{-1})\{(x_1-x_2)+(x_3-x_4)+\ldots+(x_{2n-1}-x_{2n})\}$

The statistic U is identical with that defined in (1). In the expansion of U^4, the expectation of a typical term is given by

$$E(x_{2i-1}-x_{2i})(x_{2i'-1}-x_{2i'})(x_{2i''-1}-x_{2i''})(x_{2i'''-1}-x_{2i'''})$$

$$= \Delta''\delta(2\overline{i-i'})\,\Delta''\delta(2\overline{i''-i'''}) + \Delta''\delta(2\overline{i-i''})\,\Delta''\delta(2\overline{i'-i'''}) + \Delta''\delta(2\overline{i-i'''})\,\Delta''\delta(2\overline{i''-i'})$$

$$\equiv U_{ii'i''i'''}, \quad \text{say.} \quad (44)$$

The value of this depends only on the configuration of i, i', i'', i'''. If these are placed in increasing order of magnitude, we may denote the central interval by q, and the other two by p and r, where $p \leqslant r$. The number of terms giving the same values of p, q and r, and hence the same expectations, is equal to $(n-p-q-r)$ multiplied by the factor a_{pqr} given below:

Range of p, q, r	a_{pqr}
$0<p<r;\ q>0$	$2.4!$
$0<p=r;\ q>0$	$4!$
$0=p<r;\ q>0$	$2.4!/2!$
$0=p=r;\ q>0$	$4!/(2!)^2$
$0<p<r;\ q=0$	$2.4!/2!$
$0<p=r;\ q=0$	$4!/2!$
$0=p<r;\ q=0$	$2.4!/3!$
$0=p=q=r$	1

If we write the total of these terms in each case as $(n-p-q-r)\,B_{pqr}$, we obtain the following result:

$$\begin{aligned} E(U^4) &= n^{-4} \sum_{i,i',i'',i'''=1}^{n} u_{ii'i''i'''} \\ &= n^{-4} \sum_{\substack{p,q,r \\ (p\leqslant r,\, p+q+r\leqslant n-1)}} (n-p-q-r)\,B_{pqr} \\ &= n^{-4} \sum_{r=0}^{n-1} \sum_{p=0}^{\min(r,\,n-r-1)} \sum_{q=0}^{n-p-r} (n-p-q-r)\,B_{pqr}. \end{aligned} \tag{45}$$

If for $s \geqslant s_0$, $\delta(2s)$ differs from linearity over the interval $(2s-1, 2s+1)$ by less than some small quantity ϵ, we have

$$|\Delta''\delta(2s)| < 4\epsilon. \tag{46}$$

Since

$$B_{pqr} \propto \Delta''\delta(2p)\,\Delta''\delta(2r) + \Delta''\delta(2\overline{p+q})\,\Delta''\delta(2\overline{q+r}) + \Delta''\delta(2\overline{p+q+r})\,\Delta''(2q), \tag{47}$$

B_{pqr} will be of magnitude $O(1)$ if both p and r are less than s_0, of magnitude $O(\epsilon)$ if just one of p, q, r is less than s_0, and of magnitude $O(\epsilon^2)$ if none is less than s_0.

The summation in (45) may be separated into parts having different orders of magnitude. If $s_0 \ll n$,

$$\begin{aligned} E(U^4) = n^{-4} &\sum_{r=0}^{s_0-1} \sum_{p=0}^{r} \sum_{q=0}^{n-r-p-1} (n-p-q-r)\,B_{pqr} \\ &+ n^{-4} \sum_{r=s_0}^{n-1} \sum_{p=0}^{s_0-1} \sum_{q=0}^{n-r-p-1} (n-p-q-r)\,B_{pqr} \\ &+ n^{-4} \sum_{r=s_0}^{n-1} \sum_{p=s_0}^{\min(r,\,n-r-1)} \sum_{q=0}^{\min(s_0-1,\,n-r-p-1)} (n-p-q-r)\,B_{pqr} \\ &+ n^{-4} \sum_{r=s_0}^{n-1} \sum_{p=s_0}^{\min(r,\,n-r-1)} \sum_{q=\min(s_0-1,\,n-r-p-1)}^{n-r-p-1} (n-p-q-r)\,B_{pqr}. \end{aligned} \tag{48}$$

The four terms on the right-hand side of (48) are of magnitudes respectively

$$O(n^{-2}\epsilon^0), \quad O(n^{-1}\epsilon^1), \quad O(n^{-1}\epsilon^1) \quad \text{and} \quad O(n^0\epsilon^2).$$

Hence if

$$\epsilon = o(n^{-1}) \tag{49}$$

the right-hand side of (48) will be dominated by the first term. To order n^{-2}, this may be put in a more readily computable form:

$$E(U^4) \sim n^{-2} \sum_{r=0}^{s_0-1} \sum_{p=0}^{r} \tfrac{1}{2} a_{p1r}\,\Delta''\delta(2p)\,\Delta''\delta(2r), \tag{50}$$

since

$$\sum_{q=1}^{n-r-p-1} (n-p-q-r) = \tfrac{1}{2}(n-r-p)(n-r-p-1) \sim \tfrac{1}{2}n^2, \tag{51}$$

and $a_{pqr} = a_{p1r}$ for $q > 0$. To this order, we may neglect the term in a_{p0r}.

This example has been chosen because of its comparative simplicity, to illustrate the way in which the summations in these problems have to be arranged in order to take the important terms first. It is a fourth moment, corresponding to $r = 4$ in (30), and we have verified that it is of order $n^{-4r} = n^{-2}$, provided that linearity in $\delta(s)$ is attained rapidly enough. It might be required in allowing for the effect of terms of order n^{-2} in the test of

significance of the treatment effect in (5) (cf. Small, 1954), or in obtaining control chart limits correct to order n^{-1} for the difference between means of interpenetrating systematic samples from, say, a conveyor belt.

REFERENCES

BARTLETT, M. S. (1946). Some aspects of the time-correlation problem in regard to tests of significance. *J.R. Statist. Soc.* **98**, 536.

JOWETT, G. H. (1952). The accuracy of systematic sampling from conveyor belts. *Appl. Statist.* **1**, 50.

JOWETT, G. H. (1953). The comparison of means for industrial time series. Paper given to Royal Statistical Society conference.

JOWETT, G. H. (1955*a*). Least squares regression analysis for trend-reduced time series. *J.R. Statist. Soc.* B (in the Press).

JOWETT, G. H. (1955*b*). The comparison of means of industrial time series. *Appl. Statist.* **4** (in the Press).

SMALL, V. J. (1954). M.Sc. Thesis for the University of Sheffield.

MODELS FOR TWO-DIMENSIONAL STATIONARY STOCHASTIC PROCESSES

By V. HEINE

*Applied Mathematics Laboratory, Department of Scientific and Industrial Research, Wellington, N.Z.**

To aid the analysis of two-dimensional stationary processes, three different models are considered, derived from the second-order stochastic partial differential equation. Their correlation functions are calculated, for fitting data in the form of a correlogram to the model. The corresponding Green functions describe the physical nature of the models, and in particular distinguish between space-like and time-like axes.

1. Introduction

The occurrence of stationary stochastic processes in two dimensions is well known. In analysing such data, it is at least useful, and occasionally of theoretical importance, to fit the correlogram to a particular plausible model and to estimate certain parameters in the model accordingly. However, owing to mathematical complexities, only the correlation functions generated by the very simplest models have been investigated in the past, particularly by Whittle (1954), and it would seem desirable to extend the range of available models.

We shall consider the general second-order linear stochastic partial differential equation

$$\left(a\frac{\partial^2}{\partial x^2}+2h\frac{\partial^2}{\partial x\partial y}+b\frac{\partial^2}{\partial y^2}+2g\frac{\partial}{\partial x}+2f\frac{\partial}{\partial y}+c\right)\xi(x,y)=\epsilon(x,y), \tag{1·1}$$

where ξ and ϵ are the variate and the random impulses effecting it respectively, both with means zero. This leads to three types of model, corresponding to parabolic, elliptic and hyperbolic forms, whose correlation and Green functions we shall derive.

In two-dimensional processes, one is necessarily concerned with the difference between time-like and space-like axes. In a time series, the variate can only be influenced by past events; and accordingly, a time-like axis is one, such that a random impulse can only produce an effect in one direction along it. Besides time itself, examples are the distance down a steep hillside, and the distance in the direction of a strong wind scattering seeds. Along a space-like axis, the variate depends on events in both directions. We discuss these features for each model in terms of the Green function, which represents the effect at one point, of a random impulse at another point. In fact, the whole nature of the process is best visualized by considering the Green function, which thus shows what kind of physical processes the models may represent.

2. Preliminary theory

The results of §§ 2 and 3 are analogous to those relating to one-dimensional time series, and we present them here but briefly in the particular form required later (cf. Bartlett (1946) and Daniell (1946)).

Consider the stochastic partial differential equation

$$L\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y}\right)\xi(x,y)=\epsilon(x,y). \tag{2·1}$$

* Now at Sidney Sussex College, Cambridge.

Assuming the validity of inverting the order of differentiation and integration, the solution may be written formally

$$\xi(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-u, y-v)\,\epsilon(u,v)\,du\,dv, \tag{2·2}$$

where the Green function $G(x,y)$ satisfies

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) G(x,y) = \delta(x)\,\delta(y), \tag{2·3}$$

using the Dirac delta function (cf. Van der Pol & Bremner, pp. 75, 315). The physical interpretation of (2·2) and (2·3) is that $G(x-u, y-v)$ represents the effect at the point (x,y) of a unit impulse at the point (u,v).

If the $\epsilon(u,v)$ are entirely uncorrelated random impulses, we have for the covariance functions

$$\begin{aligned}\mathrm{cov}_{\epsilon}(x,y) &= \text{expectation of } [\epsilon(u,v)\,\epsilon(u-x, v-y)] \\ &= \sigma^2\delta(x)\,\delta(y),\end{aligned} \tag{2·4}$$

and

$$\mathrm{cov}_{\xi}(x,y) = \sigma^2\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(u,v)\,G(u-x, v-y)\,du\,dv \tag{2·5}$$

by using (2·2) and (2·4). In what follows, it will be convenient to calculate the function

$$R(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(u,v)\,G(u-x, v-y)\,du\,dv, \tag{2·6}$$

which can then be normalized right at the end, to give the correlation function of ξ in the form

$$\rho(x,y) = R(x,y)/R(0,0). \tag{2·7}$$

Further, G and R must tend to zero at infinity, and R must be finite everywhere.

3. Use of the Laplace transform

The manipulative complexities, encountered in calculating the Green and correlation functions from (2·3), (2·6) and (2·7), are handled using the two-sided Laplace transform. We follow completely the notation and dictionary of formulae in Van der Pol & Bremner (1950), reference to whom will be made simply by the page numbers. Thus if $f(p,q)$ is the transform of $h(x,y)$ we write (pp. 18, 334–6)

$$h(x,y) \doteqdot f(p,q),$$

where

$$f(p,q) = pq\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-px-qy}\,h(x,y)\,dx\,dy.$$

If $g(p,q)$ be written for the transform of $G(x,y)$, (2·3) gives (pp. 48, 345, 384)

$$L(p,q)\,g(p,q) = pq,$$

i.e.

$$G(x,y) \doteqdot g(p,q) = \frac{pq}{L(p,q)}. \tag{3·1}$$

The transform of (2·6) may be written down using (3·1) and the composition product formula (pp. 39, 382); thus

$$R(x,y) \doteqdot \frac{g(p,q)\,g(-p,-q)}{pq} = \frac{pq}{L(p,q)\,L(-p,-q)}. \tag{3·2}$$

4. Standard forms

By the simple substitutions $x' = -x$, kx, $\pm y$, $x\cos\theta + y\sin\theta$ and in the hyperbolic case by $x' = x + c_1 y, y' = x + c_2 y$; the general form (1·1) can always be reduced to one of the following standard forms:

parabolic
$$\left(\frac{\partial}{\partial x}+\alpha\right)^2 - \gamma^2\left(\frac{\partial}{\partial y}\pm\beta\right); \tag{4·1}$$

elliptic
$$\left(\frac{\partial}{\partial x}-\alpha\right)^2 + \frac{\partial^2}{\partial y^2} \pm \gamma^2; \tag{4·2}$$

hyperbolic
$$\left(\frac{\partial}{\partial x}+\alpha\right)\left(\frac{\partial}{\partial y}+\beta\right) \pm \gamma^2; \tag{4·3}$$

degenerate
$$\gamma\frac{\partial^2}{\partial y^2} + \alpha\frac{\partial}{\partial y} \pm \beta; \tag{4·4}$$

where α, β, γ are all real and positive, or zero.

5. The parabolic form

Consider
$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x}+\alpha\right)^2 - \gamma^2\left(\frac{\partial}{\partial y}+\beta\right). \tag{5·1}$$

From (3·1)
$$G(x,y) \doteqdot \frac{pq}{(p+\alpha)^2 - \gamma^2(q+\beta)}, \tag{5·2}$$

hence
$$e^{\alpha x+\beta y}\,G(x,y) \doteqdot \frac{pq}{p^2-\gamma^2 q} \quad \text{(p. 374)}$$
$$\doteqdot \frac{-p}{\gamma^2}\exp\left(\frac{p^2 y}{\gamma^2}\right) U(y) \quad \text{(p. 26)},$$

hence
$$G(x,y) = \frac{-1}{2\gamma\sqrt{(\pi y)}}\exp\left(-\alpha x - \beta y - \frac{x^2\gamma^2}{4y}\right) U(y) \quad \text{(p. 21)}, \tag{5·3}$$

where
$$U(y) = 1 \quad (y \geqslant 0); \tag{5·4}$$
$$= 0 \quad (y < 0).$$

The case of the lower sign in (4·1) is covered by considering β as negative in (5·1), which makes G in (5·3) non-vanishing at infinity, and is thus inadmissible.

Equation (5·3) shows that an impulse at the origin only has an effect at positive values of y. Thus the y-axis is a time-like axis, whereas the x-axis is space-like. The Green function has the shape of a Gaussian error curve in the x-direction, with ever-increasing variance and decreasing amplitude as y increases. Hence the y-axis may well represent distance downhill, downstream or along the direction of a wind in any kind of diffusion phenomenon. Otherwise it may represent time, during which something spreads out along a line, for instance, the descendants of a plant.

From (3·2),

$$\gamma^4 R(x,y) \doteqdot \frac{pq}{\left(\frac{\gamma^2\beta-\alpha^2-p^2}{\gamma^2}\right)^2 - \left(q - \frac{2\alpha p}{\gamma^2}\right)^2}$$
$$\doteqdot \tfrac{1}{2}\exp\left(\frac{2\alpha p}{\gamma^2}y\right)\frac{\gamma^2 p}{(\gamma^2\beta-\alpha^2-p^2)}\exp\left[\frac{(p^2+\alpha^2-\gamma^2\beta)}{\gamma^2}|\gamma|\right] \quad \text{(p. 27)}. \tag{5·5}$$

Without loss of generality, consider the two quadrants with $y > 0$. Thus rearranging,

$$R(x,y) \doteqdot \frac{1}{2\gamma^2} e^{-\beta y} \left[\frac{p}{p+\alpha}(p+\alpha)\exp\left[\frac{(p+\alpha)^2 y}{\gamma^2}\right]\right]\left[\frac{p}{(\gamma^2\beta-\alpha^2)-p^2}\right]\frac{1}{p}. \tag{5·6}$$

Now
$$\frac{p}{p+\alpha}(p+\alpha)\exp\left[\frac{(p+\alpha)^2 y}{\gamma^2}\right] \doteqdot \frac{e^{-\alpha x}}{2\gamma\sqrt{(\pi y)}}\exp\left(-\frac{x^2\gamma^2}{4y}\right) \quad \text{(p. 21)}, \tag{5·7}$$

and
$$\frac{p}{(\gamma^2\beta-\alpha^2)-p^2} \doteqdot \frac{1}{2\sqrt{(\gamma^2\beta-\alpha^2)}}\exp\left(-\sqrt{(\gamma^2\beta-\alpha^2)}\,|x|\right) \quad \text{(p. 27)}. \tag{5·8}$$

We can now use the composition product rule (p. 39), also (5·7) and (5·8), in (5·6):

$$R(x,y) = (\text{constant})\frac{e^{-\beta y}}{\sqrt{y}}\int_{-\infty}^{\infty}\exp\left(-\alpha\tau - \frac{\gamma^2\tau^2}{4y} - \sqrt{(\gamma^2\beta-\alpha^2)}\,|x-\tau|\right)d\tau. \tag{5·9}$$

The integral in (5·9) may be simplified by breaking it up into two integrals over the ranges $\tau < x$ and $\tau > x$ respectively, and then using linear substitutions

$$t = \frac{\gamma}{2\sqrt{y}}\left(\tau + \frac{2\alpha y}{\gamma^2}\right) \mp \sqrt{\left\{y\left(\beta - \frac{\alpha^2}{\gamma^2}\right)\right\}}.$$

Using (2·7) to insert the appropriate normalizing constant, we finally have

$$\begin{aligned}\rho(x,y) &= \rho(-x,-y), \quad \text{where} \quad y>0,\\ &= \frac{e^{-2AB}}{\sqrt{\pi}}\int_{-\infty}^{A-B} e^{-t^2}dt\\ &\quad + \frac{e^{+2AB}}{\sqrt{\pi}}\int_{A+B}^{\infty} e^{-t^2}dt,\end{aligned} \tag{5·10}$$

where
$$A = \frac{\gamma}{2\sqrt{y}}\left(x + \frac{2\alpha y}{\gamma^2}\right), \quad B = \sqrt{\left\{y\left(\beta - \frac{\alpha^2}{\gamma^2}\right)\right\}},$$

with
$$0 \leqslant \alpha^2 < \beta\gamma^2 \quad \text{[see (5·13)]}.$$

As emphasized by Van der Pol & Bremner (1950, pp. 19, 27), to ensure the validity of the above argument, it is necessary to show that (5·2)–(5·9) possess a common non-vanishing region of convergence. This requirement eliminates several other solutions of (5·2) and (5·5) which would otherwise appear to be formally possible, and also imposes some sufficient conditions on α, β and γ. By considering (5·1), (2·3) and (2·6) directly, without the use of the Laplace transform, these conditions may be shown to be also necessary.

Thus (5·3) requires (pp. 26, 374)

$$\operatorname{Re}\left[-\beta + \left(\frac{p+\alpha}{\gamma}\right)^2\right] < \operatorname{Re} q < \infty. \tag{5·11}$$

(5·5) in addition requires (p. 378)

$$-\infty < \operatorname{Re} q < -\operatorname{Re}\left[-\beta + \left(\frac{p-\alpha}{\gamma}\right)^2\right]. \tag{5·12}$$

For these regions to overlap, we must have the left-hand expression in (5·11) less than the right-hand expression in (5·12), which leads to

$$0 \leqslant \alpha^2 < \beta\gamma^2. \tag{5·13}$$

It may be shown that if and only if (5·13) holds, do (5·2) to (5·9) have a non-vanishing common region of convergence.

6. Elliptic forms

The special form

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} - \gamma^2 \tag{6·1}$$

has been treated by Whittle (1954), who used it in discussing the yields of orange trees in a square array. For the sake of completeness, we repeat

$$G(x, y) = \frac{1}{2\pi} K_0(\gamma r) \doteqdot \frac{pq}{p^2 + q^2 - \gamma^2} \quad \text{(pp. 357, 407).} \tag{6·2}$$

$R(x, y)$ is obtained by differentiating (6·2) with respect to γ (p. 373). Then from (2·7),

$$\rho(x, y) = (\gamma r) K_1(\gamma r). \tag{6·3}$$

Here and elsewhere, $r = \sqrt{(x^2 + y^2)}$, and the K's are the modified Bessel functions of the second kind. Both the x- and y-axes are space-like, and for (6·1) the Green and correlation functions decrease monotonically equally in all directions. A simple change of scale in the x- or y-direction introduces unequal degrees of correlation along these axes. This model would be applicable to a wide variety of circumstances, such as field trials on flat land.

A form involving a further parameter is

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x} - \alpha\right)^2 + \frac{\partial^2}{\partial y^2} - \gamma^2. \tag{6·4}$$

From (3·1), (6·2) and pp. 357, 339,

$$G(x, y) = \frac{e^{\alpha x}}{2\pi} K_0(\gamma r) \doteqdot \frac{pq}{(p - \alpha)^2 + q^2 - \gamma^2}. \tag{6·5}$$

This Green function is the same as (6·2), except that the term $\exp(\alpha x)$ increases the influence an impulse has in the positive x-direction at the expense of the negative x-direction, and thus may represent the effect of a wind or a slight slope of the ground. From (3·2),

$$R(x, y) \doteqdot \frac{1}{4\alpha p}\left[\frac{pq}{(p - \alpha)^2 + q^2 - \gamma^2} - \frac{pq}{(p + \alpha)^2 + q^2 - \gamma^2}\right].$$

Using (6·5), the integration rule (p. 51) and (2·7)

$$\rho(x, y) = \frac{\sqrt{(\gamma^2 - \alpha^2)}}{\sin^{-1}(\alpha/\gamma)} \int_x^\infty \sinh(\alpha\tau) K_0(\gamma \sqrt{(\tau^2 + y^2)})\, d\tau, \tag{6·6}$$

where the normalizing constant follows from a formula in Watson (1944) on p. 388. Convergence of (6·6) requires $\alpha < \gamma$.

All other elliptic forms are inadmissible. Consider

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x} - \alpha\right)^2 + \frac{\partial^2}{\partial y^2} + \gamma^2. \tag{6·7}$$

If $\alpha = \gamma = 0$, the Green function is $\log r$ which does not vanish at infinity. If $\alpha = 0$, $\gamma \neq 0$, the Green function is $\frac{1}{4} Y_0(\gamma r)$ (p. 357), and the integral in (2·6) does not converge. If $\alpha \neq 0$, $\gamma \neq 0$, the Green function $\frac{1}{2} \exp(\alpha x) Y_0(\gamma r)$ does not vanish at infinity.

7. HYPERBOLIC FORMS

Consider
$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x}+\alpha\right)\left(\frac{\partial}{\partial y}+\beta\right)+\gamma^2. \tag{7·1}$$

From (3·1) and p. 346, and using the function U defined in (5·4),

$$G(x,y) = e^{-\alpha x-\beta y} J_0(2\gamma\sqrt{(xy)})\, U(x)\, U(y)$$
$$\doteqdot \frac{pq}{(p+\alpha)(q+\beta)+\gamma^2} \tag{7·2}$$

is the only Green function that vanishes at infinity. Its form shows that any impulse can only have an influence in the positive x- and y-directions, which therefore represent two time-like axes.

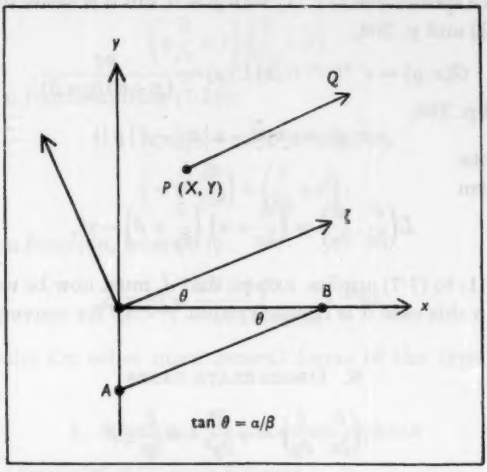


Fig. 1. Path of integration for (7·7).

From (3·2),
$$R(x,y) \doteqdot \frac{-1}{2(\beta p+\alpha q)}\left[\frac{pq}{(p+\alpha)(q+\beta)+\gamma^2} - \frac{pq}{(p-\alpha)(q-\beta)+\gamma^2}\right], \tag{7·3}$$

with region of convergence $|\operatorname{Re} p| < \alpha$, $|\operatorname{Re} q| < \beta$. If we put

$$f(x,y) = e^{-\alpha x-\beta y} J_0(2\gamma\sqrt{(xy)})\, U(x)\, U(y) - e^{\alpha x+\beta y} J_0(2\gamma\sqrt{(xy)})\, U(-x)\, U(-y)$$
$$\doteqdot \left[\frac{pq}{(p+\alpha)(q+\beta)+\gamma^2} - \frac{pq}{(p-\alpha)(q-\beta)+\gamma^2}\right], \tag{7·4}$$

then (7·3) becomes
$$2\left(\beta\frac{\partial}{\partial x}+\alpha\frac{\partial}{\partial y}\right) R(x,y) = -f(x,y). \tag{7·5}$$

Hence $R(X, Y)$ is obtained by integrating $f(x,y)$ of (7·4) in the ζ direction, i.e. along PQ (Fig. 1), making an angle $\tan^{-1}(\alpha/\beta)$ with the x-axis:

$$R(X,Y) = \frac{1}{2\sqrt{(\alpha^2+\beta^2)}}\int_{P(X,Y)}^{\infty} f(x,y)\, d\zeta. \tag{7·6}$$

In the first quadrant, $f(x, y)$ is equal to the first term in (7·4), and

$$\rho(X, Y) = \frac{2\alpha\beta}{\sqrt{(\alpha^2+\beta^2)}} \sqrt{\left(1+\frac{\gamma^2}{\alpha\beta}\right)} \int_{P(X,Y)}^{\infty} e^{-\alpha x-\beta y} J_0(2\gamma\sqrt{(xy)})\, d\zeta, \tag{7·7}$$

where the normalizing constant is obtained from a formula on p. 384 in Watson (1944). In the fourth quadrant, $f(x, y)$ is zero, so that ρ is constant along any line AB (Fig. 1), and is equal to $\rho(B)$ as given by (7·7). In the third and fourth quadrants, ρ is obtained from the relation $\rho(X, Y) = \rho(-X, -Y)$. It is necessary for the convergence of (7·7) that $\alpha > 0$ and $\beta > 0$. In evaluating (7·7) numerically, it would be helpful to employ a change of scale to make α and β each $1/\sqrt{2}$.

The form
$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x}+\alpha\right)\left(\frac{\partial}{\partial y}+\beta\right) \tag{7·8}$$

may be considered as a special case of (7·1) with $\gamma = 0$, but it is easier to start again from first principles. From (3·1) and p. 386,

$$G(x, y) = e^{-\alpha x-\beta y}\, U(x)\, U(y) \fallingdotseq \frac{pq}{(p+\alpha)(q+\beta)}. \tag{7·9}$$

From (3·2), (2·7) and p. 386,

$$\rho(x, y) = \exp(-\alpha|x|-\beta|y|) \tag{7·10}$$

over all four quadrants.

Considering the form

$$L\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \left(\frac{\partial}{\partial x}+\alpha\right)\left(\frac{\partial}{\partial y}+\beta\right) - \gamma^2, \tag{7·11}$$

the above analysis (7·1) to (7·7) applies, except that J_0 must now be replaced by I_0, and the sign of γ^2 changed. In this case it is necessary that $\gamma^2 < \alpha\beta$ for convergence.

8. Degenerate cases

Degenerate forms
$$L_1\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) = \gamma\frac{\partial^2}{\partial y^2} + \alpha\frac{\partial}{\partial y} \pm \beta \tag{8·1}$$

are strictly speaking inadmissible. For from (3·2),

$$R(x, y) = \delta(x)\, R_1(y), \tag{8·2}$$

where $R_1(y)$ is some function of y; and thus $R(x, y)$ cannot be normalized in the sense of (2·7), because of the delta function.

However, a case may well arise in practice, where the correlogram is approximately zero everywhere except along one direction, say the y-axis. It is therefore necessary to consider forms arbitrarily close to (8·1); or what is the same thing, to consider (8·1) as the result of some limiting process.

The important fact is that the result depends on the particular limiting process used. How this is possible may be seen as follows. Consider the rather restricted form

$$L\left(\eta\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) \quad \text{with limit (8·1) as } \eta \to 0. \tag{8·3}$$

The correlation function may always be written in the form

$$\rho(x, y) = \rho(0, y)\, N(x/\eta, y), \tag{8·4}$$

where $N(x/\eta, y) = 1$ for $x = 0$,

$= \text{small}$ for $x \gg \eta$, since $\rho \to 0$ as $x/\eta \to \infty$.

In the limit, the form (8·1) has the correlation function

$$\rho(x, y) = \rho_1(y)\, N(x), \tag{8·5}$$

where the null function

$$N(x) = 1 \quad (x = 0);$$
$$= 0 \quad (x \neq 0).$$

As there are many different forms satisfying (8·3), so there are different correlation functions (8·4) and (8·5). We give the following example.

Limiting form: $\dfrac{\partial}{\partial y} + \beta.$

I form: $\left(\eta \dfrac{\partial}{\partial x} + 1\right)\left(\dfrac{\partial}{\partial y} + \beta\right).$

Limiting correlation function, from (7·10):

$$\rho(x, y) = e^{-\beta|y|} N(x).$$

II form: $\left(-\eta^2 \dfrac{\partial^2}{\partial x^2}\right) + \left(\dfrac{\partial}{\partial y} + \beta\right).$

Limiting correlation function, from (5·1):

$$\rho(x, y) = \frac{2N(x)}{\sqrt{\pi}} \int_{\sqrt{(\beta|y|)}}^{\infty} e^{-t^2} dt.$$

Corresponding results for other more general forms of the type (8·1) may be easily obtained.

9. SUMMARY OF ALLOWED MODELS

The allowed standard forms (cf. § 4) are as follows:

Parabolic; x-axis space-like, y-axis time-like:

$$\left(\frac{\partial}{\partial x} + \alpha\right)^2 - \gamma^2\left(\frac{\partial}{\partial y} + \beta\right),$$
$$0 \leqslant \alpha^2 < \beta\gamma^2;$$

Green function (5·3), correlation function (5·10):

Elliptic; x- and y-axes space-like:

$$\left(\frac{\partial}{\partial x} - \alpha\right) + \frac{\partial^2}{\partial y^2} - \gamma^2,$$
$$0 \leqslant \alpha < \gamma;$$

Green function (6·5), correlation function (6·6):

Hyperbolic; x- and y-axes time-like:

$$\left(\frac{\partial}{\partial x} + \alpha\right)\left(\frac{\partial}{\partial y} + \beta\right) + \gamma^2,$$
$$\alpha > 0, \quad \beta > 0, \quad \gamma^2 \geqslant 0;$$

Green function (7·2), correlation function (7·7): also

$$\left(\frac{\partial}{\partial x}+\alpha\right)\left(\frac{\partial}{\partial y}+\beta\right)-\gamma^2,$$

$$\alpha>0, \quad \beta>0, \quad 0\leqslant\gamma^2<\alpha\beta;$$

for Green and correlation functions, see above (7·11).

The problem was suggested by Dr P. Whittle, to whom I am also indebted for helpful discussion.

REFERENCES

BARTLETT, M. S. (1946). *Stochastic Processes*. Mimeographed, North Carolina lecture notes.
DANIELL, P. J. (1946). Contribution to discussion on stochastic processes. *J. R. Statist. Soc.* B, **8**, 88.
VAN DER POL, B. & BREMNER, H. (1950). *Operational Calculus based on the Two-sided Laplace Transform*. Cambridge University Press.
WATSON, G. N. (1944). *The Theory of Bessel Functions*, 2nd ed. Cambridge University Press.
WHITTLE, P. (1954). On stationary processes in the plane. *Biometrika*, **41**, 434.

SOME PROBLEMS IN THE THEORY OF PROVISIONING AND OF DAMS

By J. GANI

Australian National University, Canberra, A.C.T.

The paper begins with a review of two problems in the theory of provisioning with a discrete stock considered by Pitt (1946), and a short account of Moran's work (1954) in the theory of finite dams. It is pointed out that these provide two different methods of attack, each appropriate to certain conditions, on problems in the probability theory of a general storage function $S(t)$ defined at time t by

$$S(t) = I(t) - D(t) - F(t),$$

where $I(t)$, $D(t)$, $F(t)$ are respectively an input, output and overflow function. This storage function is identified with the stock deficit in provisioning theory, or the dam content in dam theory, so that any problem and its solution in the one theory has an exact analogue in the other.

The paper continues with the use of Pitt's results of the theory of provisioning in two analogous cases of the infinite discrete dam. This is followed by the application of Moran's methods of the theory of dams in some analogous problems of provisioning with a discrete finite stock; exact solutions are obtained for the discrete and continuous cases of a particular problem in which ordering and replacement times coincide.

The paper closes with an exact solution of the general storage problem in the case of a finite continuous storage function $S(t)$, fed by a discrete input function of Poisson type, with a continuous output which has a steady rate when $S(t) \neq 0$, and is zero if $S(t) = 0$, and an overflow function such that $S(t)$ never exceeds a prescribed value.

I. Problems in the theory of provisioning

In his work on the theory of provisioning, Pitt (1946) is concerned with the derivation of stationary probability distributions for a discrete stock function $s(t)$. This stock function represents the number of components in stock at a time t, and is allowed to take integral values ranging between K and $-\infty$, where $K = s(0)$ is the initial stock, and negative values indicate that components may be borrowed if they are not in stock. The function is defined for all values of time t (t continuous) by the equation

$$s(t) = K + r(t) - c(t), \tag{1}$$

where $c(t)$ is the consumption function, the number of components consumed up to time t, a random increasing step function taking positive integral values, and $r(t)$ is the replacement function, the number of components delivered and added to the stock up to time t, a function depending on $c(t)$ and always smaller than or equal to it.

The probability distribution associated with the consumption function $c(t)$ is the Poisson with parameter a, so that the probabilities that in a small interval of time δt there be one and no components consumed respectively are

$$\left.\begin{aligned} \Pr\{c(t+\delta t) - c(t) = 1\} &= a\,\delta t, \\ \Pr\{c(t+\delta t) - c(t) = 0\} &= 1 - a\,\delta t, \end{aligned}\right\} \tag{2}$$

and increases in $c(t)$ in non-overlapping intervals are independent. Two types of replacement functions are considered:

Type 1. $$r_1(t) = [c(t-T)\,M^{-1}]\,M, \tag{3}$$

where we define $[x]$ as the integral part of x; here, orders for a constant number of components M are sent out at irregular time intervals $t - T$, when $c(t-T)$ is an integral multiple of M, and are delivered and added to the stock at time t, after a positive time-lag T;

Type 2. $$r_2(t) = c\{[(t-T)aM^{-1}]Ma^{-1}\}; \tag{4}$$

here, orders for an irregular number of components equal to the consumption in the previous time interval Ma^{-1} are sent out at regular times kMa^{-1} ($k = 1, 2, \ldots$), and are delivered and added to stock at times $kMa^{-1}+T$ after a positive time-lag T.

For both these types of replacement function, the stationary probability distributions of $s(t)$ are derived by arguments which will be followed exactly in § IV of the paper. Using these probability distributions, Pitt concludes with a comparison of the means of both the positive and negative values of the stock function $s(t)$ when the replacement functions are of types 1 and 2.

II. Problems in the theory of dams

The problem discussed by Moran (1954), that of obtaining the stationary probability distribution of the amount of water in a dam, appears at first sight to be of a different nature. A finite dam of capacity K units, whose content at the times $t = 0, 1, 2, \ldots$, after water has been released according to a prescribed rule, is $Z(t)$, where $Z(t)$ will be found to lie in the range $(0, K-M)$, is subject to the following conditions:

(1) a discrete input $X(t)$ which flows into the dam during the interval of time $(t, t+1)$, where the series $\{X(t)\}$ is serially independent, and the probability that the input be i units ($i = 0, 1, 2, \ldots$) is p_i;

(2) an overflow rule such that in any time interval $(t, t+1)$ there is no overflow from the dam if $Z(t)+X(t) \leqslant K$, and there is an overflow $Z(t)+X(t)-K$ if $Z(t)+X(t) > K$;

(3) a release rule such that at time $t+1$, M units are released from the dam if its total content $Z(t)+X(t) \geqslant M$, and the total content $Z(t)+X(t)$ is released if this is less than M.

Then, provided $K-M \geqslant M$, a system of equations is derived relating the probabilities $P_0, P_1, \ldots, P_{K-M}$ that $Z(t)$ be equal to $0, 1, \ldots, K-M$ at time t, with the probabilities $P_0^{(1)}, P_1^{(1)}, \ldots, P_{K-M}^{(1)}$ that $Z(t+1)$ be equal to $0, 1, \ldots, K-M$ at time $t+1$. These are:

$$\left.\begin{aligned}
P_0^{(1)} &= P_0(p_0+p_1+\ldots+p_M)+P_1(p_0+\ldots+p_{M-1})+\ldots+P_M p_0,\\
P_1^{(1)} &= P_0 p_{M+1} \qquad\quad +P_1 p_M \qquad\quad +\ldots+P_{M+1}p_0,\\
&\cdots\cdots\cdots

III. PROVISIONING AND DAM PROBLEMS AS GENERAL STORAGE PROBLEMS

We now proceed to identify these problems of the stock function $s(t)$ of provisioning, and of the dam content $Z(t)$ in dam theory as particular problems of a general storage function $S(t)$. This may be discrete or continuous, and defined for continuous time t or at fixed times such as $t = 0, 1, 2, \ldots$, according to the conditions of the problem, by the equation

$$S(t) = I(t) - D(t) - F(t), \tag{6}$$

where $I(t)$, $D(t)$, $F(t)$ are positive increasing functions, discrete or continuous. $I(t)$ is a random input function which feeds the storage function, $D(t)$ is an output function which depletes the storage function and may depend on $I(t)$ or be otherwise defined, and $F(t)$ is an overflow function which operates to restrict the range of $S(t)$ below a prescribed maximum value.

To show that the problems of Pitt's stock function $s(t)$ of provisioning can be regarded as problems of a particular storage function $S(t)$, we consider the previously defined stock function (1),

$$s(t) = K + r(t) - c(t),$$

which lies in the range $-\infty < s(t) \leqslant K$. We note that this equation holds at any time t (t continuous), and that $c(t)$, the random discrete consumption function, and $r(t)$, the discrete replacement function ($r(t) \leqslant c(t)$), are defined as in § I. We can here equally well consider the stock deficit,

$$K - s(t) = c(t) - r(t), \tag{7}$$

which ranges from zero when the stock is at its maximum K, to ∞ as the stock decreases through 0 to $-\infty$. Naturally this is possible only when the stock has a maximum value K which it cannot exceed. Comparing this with a storage function

$$S(t) = I(t) - D(t), \tag{8}$$

for which $F(t) = 0$ for all t, it is clear that a discrete storage function $S(t)$ defined in the range $(0, \infty)$ for all time t (t continuous), is identifiable with the deficit $K - s(t)$. Similarly, $I(t)$, a discrete random input function, is identifiable with the discrete random consumption function $c(t)$, and $D(t)$, a discrete output function ($D(t) \leqslant I(t)$), with a discrete replacement function $r(t)$. $F(t)$ will be zero at all times, since the range of the stock deficit is $(0, \infty)$, and no overflow is required to restrict this range below any prescribed value.

In the specific problems considered by Pitt, where the consumption function $c(t)$ was of Poisson type with parameter a, and the replacement functions of the types previously defined in (3), (4), as $r_1(t)$ and $r_2(t)$, the general storage equation would be such that

$$S(t) = K - s(t), \quad I(t) = c(t), \quad D(t) = r_1(t) \quad \text{or} \quad r_2(t), \quad F(t) = 0.$$

However, generally, for any stock function $s(t)$, whether continuous or discrete, defined for continuous time t or at discrete intervals of time, providing only that the maximum stock K is finite, it is always possible to frame the general storage equation (6) for the stock deficit. The finite maximum stock K results in the storage function $S(t)$ having a range with lower bound 0; its upper bound may be infinite, as in the case just considered, or finite as we shall see in § V. $I(t)$, $D(t)$ and $F(t)$ will be appropriately chosen to fit the conditions of the problem.

Similarly, we proceed to show that Moran's discrete dam content $Z(t)$ with range $(0, K - M)$ can be regarded as a particular storage function. In this case, however, the function $Z(t)$

is defined only at the times $t = 0, 1, 2, \ldots$, and we shall also define the storage function at these times. In our storage equation (6), we may identify $S(t)$ with $Z(t) - Z(0)$, the increase in dam content after a time t; if, without loss in generality, we assume that $Z(0) = 0$, that is, the dam is empty at time $t = 0$, we have that $S(t)$ the storage function is equal to the dam content $Z(t)$. The random discrete input function $I(t)$ would be the sum of the discrete random inputs $X(0), X(1), \ldots, X(t-1)$ in the time intervals $(0, 1), (1, 2), \ldots, (t-1, t)$, so that

$$I(t) = \sum_{\tau=1}^{t} X(\tau - 1).$$

The discrete output function $D(t)$ would be the sum of discrete outputs released at times $\tau = 1, 2, \ldots, t$; let these be written $d(1), d(2), \ldots, d(t)$, then we have that

$$d(\tau) = M \quad \text{if} \quad Z(\tau-1) + X(\tau-1) \geqslant M,$$

or

$$d(\tau) = Z(\tau-1) + X(\tau-1) \quad \text{if this is less than } M,$$

so that the output function is

$$D(t) = \sum_{\tau=1}^{t} d(\tau).$$

Finally, the overflow function $F(t)$ would be the sum of the discrete overflows in the time intervals $(0, 1), \ldots (t-1, t)$; let these be written $f(0), \ldots, f(t-1)$, then we have that

$$f(\tau-1) = 0 \quad \text{if} \quad Z(\tau-1) + X(\tau-1) \leqslant K,$$

or

$$f(\tau-1) = Z(\tau-1) + X(\tau-1) - K \quad \text{if this is greater than zero,}$$

so that the overflow function is

$$F(t) = \sum_{\tau=1}^{t} f(\tau - 1).$$

The equation for $Z(t)$ at times $t = 0, 1, 2, \ldots$ could be written

$$Z(t) = \sum_{\tau=1}^{t} X(\tau-1) - \sum_{\tau=1}^{t} d(\tau) - \sum_{\tau=1}^{t} f(\tau-1), \tag{9}$$

a particular case of the storage function $S(t)$. For a continuous dam content, a similar equation could also be written. More generally, for a $Z(t)$ continuous or discrete, defined at any time t (t continuous) or at discrete times, with a finite or infinite range, it is possible to frame a general storage equation.

In neither the case of problems of provisioning nor of dam problems will the storage equation itself necessarily provide a method of solution enabling the stationary probability distribution of the function to be found. But as we are always able to reduce provisioning or dam problems to a general storage problem, this means that a provisioning problem can be interpreted equally well as a dam problem, and vice versa. In other words, any problem and its solution in the one theory will have an exact analogue in the other; equally, any method of attack on a problem in the one theory may be useful in analogous problems of the other.

We proceed to interpret Pitt's results in provisioning theory as solutions in two analogous cases of the infinite dam. We shall also apply Moran's methods in the theory of dams to some problems in the theory of provisioning.

IV. INTERPRETATION OF PITT'S RESULTS IN THE CASE OF THE INFINITE DAM

Pitt's equation for the stock deficit

$$K-s(t)=c(t)-r(t) \quad (r(t)\leqslant c(t)),$$

was seen to be an example of the storage equation

$$S(t)=I(t)-D(t),$$

where at any time t (t continuous), $S(t)$ is discrete and lies in the range $(0,\infty)$, $I(t)$ is a random discrete input function, and $D(t)$ is a discrete output function depending on $I(t)$, $(D(t)\leqslant I(t))$. Equally, a storage equation of this form can represent the discrete dam content $Z(t)$ $(S(t)=Z(t))$, at any time t, where the dam is infinite $(0\leqslant Z(t)<\infty)$. In this case, interpreting the specific conditions of Pitt's problems of provisioning for an infinite dam, we have that the discrete input function $I(t)$ will be of Poisson type with parameter a such that, in a small interval of time δt, the probabilities that one and no units of water flow into the dam are respectively

$$\left.\begin{aligned}\Pr\{I(t+\delta t)-I(t)=1\}&=a\,\delta t,\\ \Pr\{I(t+\delta t)-I(t)=0\}&=1-a\,\delta t,\end{aligned}\right\} \tag{10}$$

and increases in $I(t)$ in non-overlapping intervals are independent. The release rules corresponding to Pitt's two replacement functions are such that the resulting output functions will be:

Type 1. $$D_1(t)=[I(t-T)\,M^{-1}]\,M; \tag{11}$$

here, releases of M units are made at irregular time intervals t, after a time-lag T from the times $t-T$ when $I(t-T)$ is an integral multiple of M. In practice, this might mean that at every time when the input increased by M units, a decision would be taken to release M units from the dam, but this decision would be carried out after a time-lag T;

Type 2. $$D_2(t)=I\{[(t-T)\,aM^{-1}]\,Ma^{-1}\}; \tag{12}$$

here, releases of an irregular number of units equal to the input in the previous time interval $((k-1)\,Ma^{-1}, kMa^{-1})$ are made at regular times $kMa^{-1}+T$ $(k=1,2,\ldots)$, with time-lag T after the interval. In practice, this might mean that at the regular times kMa^{-1}, a decision would be taken to release the input in the previous interval Ma^{-1}, but this decision would be carried out after a time-lag T.

We proceed to follow Pitt exactly in deriving the stationary distributions of the dam content $Z(t)$ when the release rules of types 1 and 2 are functioning, and in comparing the mean contents of the dam for these cases.

(1) *The stationary distribution of $Z(t)$ for release rule of type* 1

In this case, where the output function is given by (11), the storage equation for the dam content can be written

$$Z(t)=I(t)-[I(t-T)\,M^{-1}]\,M \quad (0\leqslant Z(t)<\infty).$$

Let the input function at times t, and $t-T$ be such that

$$I(t-T)=i \quad \text{and} \quad I(t)-I(t-T)=j,$$

where i and j can take the values $0, 1, 2, \ldots$. Then, writing

$$P_n(t)=\Pr\{Z(t)=n\} \quad (n=0,1,2,\ldots),$$

for the probability that the dam content be n units at time t, we have that

$$P_n(t) = \Pr\{i+j-[iM^{-1}]\,M = n\}.$$

Now suppose that $\quad kM \leqslant i \leqslant (k+1)\,M-1 \quad (k=0,1,\ldots),$

then it follows that

$$[iM^{-1}] = k, \quad i = kM+v \quad \text{and} \quad j = n-v,$$

where v can take the values $0, 1, \ldots, M-1$. From the independence condition for increases in $I(t)$ in non-overlapping intervals, it follows that

$$\begin{aligned} P_n(t) &= \sum_{i,j} \Pr\{I(t-T)=i\}\,.\,\Pr\{I(t)-I(t-T)=j\} \\ &= \sum_{v=0}^{M-1} \sum_{k=0}^{\infty} e^{-a(t-T)} \frac{\{a(t-T)\}^{kM+v}}{(kM+v)!}\, g_{n-v}, \end{aligned} \tag{13}$$

where we write $\quad g_r = e^{-aT}\frac{(aT)^r}{r!} \quad$ for $\quad r \geqslant 0,$

and $\quad g_r = 0 \quad$ for $\quad r < 0.$

The stationary distribution is obtained when $t \to \infty$; we shall write it $P_n = \lim_{t\to\infty} P_n(t)$. It follows that

$$P_n = \sum_{v=0}^{M-1} g_{n-v} \lim_{t\to\infty} \sum_{k=0}^{\infty} e^{-a(t-T)} \frac{\{a(t-T)\}^{kM+v}}{(kM+v)!}. \tag{14}$$

Now consider the series $\sum_{k=0}^{\infty} e^{-y}\frac{y^{kM+v}}{(kM+v)!}$, where $y = a(t-T)$; if $y = v+(N+1)\,M$, we have that since $v \leqslant M-1$, then for $k=0$,

$$M^{-1}\left\{1+y+\ldots+\frac{y^v}{v!}\right\} < \frac{y^v}{v!} < M^{-1}\left\{\frac{y^v}{v!}+\ldots+\frac{y^{v+M-1}}{(v+M-1)!}\right\}. \tag{15}$$

For $k = 1, 2, \ldots, N$, the following inequalities hold:

$$\begin{aligned} M^{-1}\left\{\frac{y^{v+(k-1)M+1}}{(v+(k-1)\,M+1)!}+\ldots+\frac{y^{v+kM}}{(v+kM)!}\right\} \\ < \frac{y^{v+kM}}{(v+kM)!} < M^{-1}\left\{\frac{y^{v+kM}}{(v+kM)!}+\ldots+\frac{y^{v+(k+1)M-1}}{(v+(k+1)\,M-1)!}\right\}, \end{aligned} \tag{16}$$

but for $k = N+1$, only one side of the inequality is valid

$$M^{-1}\left\{\frac{y^{v+NM+1}}{(v+NM+1)!}+\ldots+\frac{y^{v+(N+1)M}}{(v+(N+1)\,M)!}\right\} < \frac{y^{v+(N+1)M}}{(v+(N+1)\,M)!}. \tag{17}$$

Further, for all values of k greater than $N+1$ $(k = N+2, \ldots)$, the inequalities are reversed

$$\begin{aligned} M^{-1}\left\{\frac{y^{v+kM+1}}{(v+kM+1)!}+\ldots+\frac{y^{v+(k+1)M}}{(v+(k+1)\,M)!}\right\} \\ < \frac{y^{v+kM}}{(v+kM)!} < M^{-1}\left\{\frac{y^{v+(k-1)M}}{(v+(k-1)\,M)!}+\ldots+\frac{y^{v+kM-1}}{(v+kM-1)!}\right\}. \end{aligned} \tag{18}$$

Summing the inequalities (15), (16), (17) and (18), for all values of k, and multiplying the result by e^{-y}, we obtain

$$M^{-1}\left\{1-\sum_{i=1}^{M} e^{-y}\frac{y^{v+(N+1)M+i}}{(v+(N+1)\,M+i)!}\right\} < \sum_{k=0}^{\infty} e^{-y}\frac{y^{v+kM}}{(v+kM)!} < M^{-1}\left\{1+e^{-y}\frac{y^{v+(N+1)M}}{(v+(N+1)\,M)!}\right\}.$$

Now let $N \to \infty$; y will also tend to infinity, and we have that

$$M^{-1} \leqslant \lim_{y \to \infty} \sum_{k=0}^{\infty} e^{-y} \frac{y^{v+kM}}{(v+kM)!} \leqslant M^{-1},$$

so that the equation (14) for P_n can be written

$$P_n = M^{-1} \sum_{v=0}^{M-1} g_{n-v}. \tag{19}$$

The mean content of the dam is given by

$$\begin{aligned} \mathscr{E}(Z) &= M^{-1} \sum_{n=0}^{\infty} n \sum_{v=0}^{M-1} g_{n-v} \\ &= M^{-1} \left\{ \sum_{n=0}^{\infty} (n + (n+1) + \ldots + (n+M-1)) \right\} g_n \\ &= aT + \tfrac{1}{2}(M-1). \end{aligned} \tag{20}$$

(2) *The stationary distribution of $Z(t)$ for release rule of type 2*

In this case, the output function is (12), and releases are made at regular times $kMa^{-1}+T$ $(k = 0, 1, \ldots)$. Consider the time interval

$$kMa^{-1} + T \leqslant t < (k+1)\, Ma^{-1} + T$$

between two releases, where T may have any value ($T \gtrless Ma^{-1}$); at any time in this interval, the content $Z(t)$ of the dam has the value

$$\begin{aligned} Z(t) &= I(t) - I([(t-T)\, aM^{-1}]\, Ma^{-1}) \\ &= I(t) - I(kMa^{-1}). \end{aligned}$$

Now if we write for the probability that the dam content be n units at time t,

$$P_n(t) = \Pr\{Z(t) = n\} \quad (n = 0, 1, \ldots),$$

then we have that

$$\begin{aligned} P_n(t) &= \Pr\{I(t) - I(kMa^{-1}) = n\} \\ &= e^{-a(t-kMa^{-1})} \frac{(at-kM)^n}{n!}. \end{aligned} \tag{21}$$

This is periodic in Ma^{-1}, so that in order to obtain the stationary distribution $\{P_n\}$, we write

$$\begin{aligned} P_n &= aM^{-1} \int_{t=kMa^{-1}+T}^{(k+1)Ma^{-1}+T} P_n(t)\, dt \\ &= M^{-1} \int_{u=0}^{M} e^{-(u+aT)} \frac{(u+aT)^n}{n!}\, du, \end{aligned}$$

where

$$u = a(t - kMa^{-1} - T).$$

On expanding the binomial, we have

$$\begin{aligned} P_n &= \sum_{v=0}^{n} e^{-aT} \frac{(aT)^{n-v}}{(n-v)!} M^{-1} \int_0^M e^{-u} \frac{u^v}{v!}\, du \\ &= \sum_{v=0}^{n} g_{n-v} Q_v, \end{aligned} \tag{22}$$

where

$$Q_v = M^{-1} \int_0^M e^{-u} \frac{u^v}{v!}\, du.$$

The mean of the dam content in this case is given by

$$\begin{aligned}\mathscr{E}(Z) &= \sum_{n=0}^{\infty} n \sum_{v=0}^{n} g_{n-v} Q_v \\ &= \sum_{v=0}^{\infty} Q_v \sum_{n=0}^{\infty} (n+v) g_n \\ &= \sum_{v=0}^{\infty} Q_v (aT+v).\end{aligned}$$

It is easy to see that

$$\sum_{v=0}^{\infty} Q_v = M^{-1} \int_0^M \sum_{v=0}^{\infty} e^{-u} \frac{u^v}{v!} du = 1;$$

$$\sum_{v=0}^{\infty} v Q_v = M^{-1} \int_0^M \sum_{v=0}^{\infty} v e^{-u} \frac{u^v}{v!} du = \tfrac{1}{2} M,$$

whence the mean is obtained as

$$\mathscr{E}(Z) = aT + \tfrac{1}{2} M. \tag{23}$$

Comparing (23) with (20), we note that the second release rule gives a larger mean content of the dam than the first.

V. Application of Moran's methods to provisioning with a finite positive stock

In dealing with provisioning, our notation will be more suggestive of the conditions of the problems if we write the general storage equation in the form

$$S(t) = K - s(t) = c(t) - r(t) - F(t), \tag{24}$$

where $S(t)$ is the stock deficit, $s(t)$ the stock with maximum value K, and $c(t)$, $r(t)$ and $F(t)$, consumption, replacement and overflow functions respectively.

Pitt considered cases where, since components not in stock could be borrowed, the deficit could become infinite, so that for replacement functions always smaller than or equal to consumption functions ($r(t) \leqslant c(t)$), no overflow function was necessary, and $F(t)$ was equal to zero. We shall discuss the case, in practice equally realistic, where the stock is finite and positive ($0 \leqslant s(t) \leqslant K$), so that the deficit is also finite and positive ($0 \leqslant S(t) \leqslant K$). This is the case where no borrowing is allowed, and consumption is lost between times when the stock becomes zero (or the deficit K) until there is a replacement which raises the stock above zero (or decreases the deficit below K). This will require, in cases where $r(t)$ is smaller than $c(t)$, that an overflow function operates so as to restrict the deficit to a value no greater than K, $S(t) \leqslant K$. In these conditions it is difficult to discuss $S(t)$ for all time t (t continuous), though it remains comparatively simple to consider it at specific time intervals. This was the case for the dam content $Z(t)$, defined at times $t = 0, 1, 2, \ldots$, for which Moran framed a set of equations which gave its stationary probability distribution; we proceed to apply Moran's methods to three particular problems of provisioning with finite positive stock.

(1) *Provisioning with replacements at fixed times $t = 1, 2, \ldots$, and no time-lag*

This problem was designed to provide an example of an exact analogue in provisioning theory of the dam problem considered by Moran; in practice, however, it gives rise to a perfectly possible situation. Suppose that at fixed times $t = 1, 2, \ldots$, a consignment of M components arrives at a store for a stock replacement; if at these times there is a deficit equal to or greater than M, the replacement consists of all M components, but if the deficit

is smaller than M, only enough components are added to the stock from the consignment to reduce the deficit to zero. There is no time-lag between the evaluation of the deficit at times $t-0$, and the deliveries of components for replacements at times t. This might well be the case for a truck containing a consignment of M components which, in its regular deliveries at a store, discharges for replacements only as many of the components as may be required. This replacement policy can be expressed by the replacement function $r(t)$, a step function with jumps at times $t = 1, 2, \ldots$, such that

$$r(t)-r(t-1) = M \quad \text{if} \quad S(t-0) \geqslant M,$$
$$\text{or} \qquad r(t)-r(t-1) = S(t-0) \quad \text{if} \quad S(t-0) < M.$$

Let the discrete consumption function $c(t)$ for the store be defined in any interval of time θ by the distribution $\{p_i(\theta)\}$ $(i = 0, 1, 2, \ldots)$, such that

$$\left.\begin{array}{l} (1)\ \Pr\{c(t+\theta)-c(t)=i\} = p_i(\theta), \\ (2)\ \text{increases in } c(t) \text{ in non-overlapping time intervals are independent.} \end{array}\right\} \tag{25}$$

The consumption will be met by components from stock only so long as the stock is positive (or the deficit less than K), but when the stock is zero (or the deficit K), no further orders can be met until a replacement arrives. This condition defines the overflow function $F(t)$, for in the interval of time $(t-1, t)$ between replacements, the overflow required to restrict the deficit to a value K will be such that

$$F(t)-F(t-1) = 0 \quad \text{if} \quad S(t-1)+c(t)-c(t-1) < K,$$
$$\text{or} \qquad F(t)-F(t-1) = S(t-1)+c(t)-c(t-1)-K \quad \text{if this is greater than zero.}$$

The storage equation (24) can now be written for times $t = 0, 1, 2, \ldots$, since the consumption, replacement and overflow functions are all defined; this, however, presents no great advantage as it does not lead to a method of solving the problem of finding the stationary probability distribution for the deficit $S(t)$.

If for $t = 0, 1, 2, \ldots$ we write the probabilities that the deficit be i components at times t and $t+1$ respectively as

$$P_i = \Pr\{S(t) = i\}, \quad P_i^{(1)} = \Pr\{S(t+1) = i\},$$

where i takes the values $0, 1, \ldots, K-M$, and if we write $p_i = p_i(1)$ and $q_j = \sum_{i=j}^{\infty} p_i$, then for $K-M \geqslant M$ the equations relating the probabilities are identical to (5), those obtained by Moran in his dam problem. In matrix form, we write $\mathbf{P}^{(1)} = \mathbf{pP}$, where $\mathbf{P}$, $\mathbf{P}^{(1)}$ are column vectors with elements P_i, $P_i^{(1)}$ respectively, and $\mathbf{p}$ is the matrix of coefficients. Exactly as in Moran's problem, the stationary distribution $P_0, P_1, \ldots, P_{K-M}$ is obtained by solving the equations $\mathbf{P} = \mathbf{pP}$ together with $\sum_{i=0}^{K-M} P_i = 1$; these $\{P_i\}$ cannot always be obtained in an explicit form, but they can always be evaluated numerically for any given values of the $\{p_i\}$.

(2) *Provisioning with replacements at fixed times* $t = kMa^{-1}+T$ $(k = 1, 2, \ldots)$, *with time-lag* $T < Ma^{-1}$

(a) *General equations for the stationary distributions*

This problem stems from Pitt's second problem of provisioning under replacement policy of type 2, with replacement function (4). The differences arising are due to the restriction of

the stock $s(t)$ to the range $0 \leqslant s(t) \leqslant K$, or the deficit $S(t)$ to the range $0 \leqslant S(t) \leqslant K$, and to the consideration of the deficit at specific time intervals only. The consumption function $c(t)$ is defined by (25) as in the previous problem, and the replacement function (4) is

$$r_2(t) = c([(t-T)aM^{-1}]\,Ma^{-1}),$$

where an irregular number of components equal to the consumption in the previous time interval Ma^{-1} are ordered at regular times kMa^{-1} ($k = 1, 2, \ldots$), and are delivered at times $kMa^{-1}+T$, after a positive time-lag T, where $T < Ma^{-1}$. As the deficit is restricted to values no greater than K, consumption is lost in the intervals after the deficit reaches the value K until a replacement arrives to reduce the deficit to a value less than K. This defines the overflow function in such a way that, in the intervals kMa^{-1}, $kMa^{-1}+T$, between orders and deliveries, the overflow is

$$\left.\begin{aligned} &F(kMa^{-1}+T)-F(kMa^{-1}) = 0 \text{ if } S(kMa^{-1})+c(kMa^{-1}+T)-c(kMa^{-1}) \leqslant K \\ \text{or}\quad &F(kMa^{-1}+T)-F(kMa^{-1}) \\ &\quad = S(kMa^{-1})+c(kMa^{-1}+T)-c(kMa^{-1})-K \text{ if this is greater than zero.} \end{aligned}\right\} \tag{26}$$

Similarly, in the intervals $kMa^{-1}+T$, $(k+1)\,Ma^{-1}$, between deliveries and new orders, the overflow is

$$F((k+1)\,Ma^{-1})-F(kMa^{-1}+T) = 0 \text{ if } S(kMa^{-1}+T)+c((k+1)\,Ma^{-1})-c(kMa^{-1}+T) \leqslant K$$

or

$$\begin{aligned} &F((k+1)\,Ma^{-1})-F(kMa^{-1}+T) \\ &\quad = S(kMa^{-1}+T)+c((k+1)\,Ma^{-1})-c(kMa^{-1}+T)-K \text{ if this is greater than zero.} \end{aligned}$$

The storage equation (24) is now fully defined at times kMa^{-1}, $kMa^{-1}+T$, but, as in the previous problem, this does not lead to a solution of the stationary distribution of $S(t)$, the deficit. We can, however, obtain sets of equations relating the probabilities

$$P_0, P_1, P_2, \ldots, P_K, \quad P_0^{(1)}, P_1^{(1)}, P_2^{(1)}, \ldots, P_K^{(1)}, \quad \text{and} \quad P_0^{(2)}, P_1^{(2)}, P_2^{(2)}, \ldots, P_K^{(2)},$$

that the deficit $S(t)$ be equal to $0, 1, 2, \ldots, K$, at the times kMa^{-1}, $kMa^{-1}+T$, $(k+1)\,Ma^{-1}$ respectively. Writing $p_i = p_i(T)$, and $q_j = \sum_{i=j}^{\infty} p_i$, we have for the times kMa^{-1}, $kMa^{-1}+T$, the relations

$$\left.\begin{aligned} P_0^{(1)} &= p_0(P_0+P_1+\ldots+P_{K-1})+q_0P_K, \\ P_1^{(1)} &= p_1(P_0+P_1+\ldots+P_{K-2})+q_1P_{K-1}, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ P_K^{(1)} &= q_KP_0, \end{aligned}\right\} \tag{27}$$

or in matrix form, $\mathbf{P}^{(1)} = \mathbf{pP}$:

$$\begin{pmatrix} P_0^{(1)} \\ P_1^{(1)} \\ P_2^{(1)} \\ \vdots \\ P_K^{(1)} \end{pmatrix} = \begin{pmatrix} p_0 & p_0 & \ldots & \ldots & \ldots & p_0 & q_0 \\ p_1 & p_1 & \ldots & \ldots & p_1 & q_1 & 0 \\ p_2 & p_2 & \ldots & p_2 & q_2 & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ q_K & 0 & \ldots & \ldots & \ldots & \ldots & 0 \end{pmatrix} \begin{pmatrix} P_0 \\ P_1 \\ P_2 \\ \vdots \\ P_K \end{pmatrix}. \tag{28}$$

Similarly, writing $p_i^{(1)} = p_i(Ma^{-1} - T)$ and $q_j^{(1)} = \sum_{i=j}^{\infty} p_i^{(1)}$, we have for the times $kMa^{-1} + T$, $(k+1)Ma^{-1}$, the relations

$$\left.\begin{aligned} P_0^{(2)} &= p_0^{(1)} P_0^{(1)}, \\ P_1^{(2)} &= p_1^{(1)} P_0^{(1)} + p_0^{(1)} P_1^{(1)}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ P_K^{(2)} &= q_K^{(1)} P_0^{(1)} + q_{K-1}^{(1)} P_1^{(1)} + \ldots + q_0^{(1)} P_K^{(1)}, \end{aligned}\right\} \tag{29}$$

or in matrix form, $\mathbf{P}^{(2)} = \mathbf{p}^{(1)}\mathbf{P}^{(1)}$:

$$\begin{pmatrix} P_0^{(2)} \\ P_1^{(2)} \\ P_2^{(2)} \\ \vdots \\ P_K^{(2)} \end{pmatrix} = \begin{pmatrix} p_0^{(1)} & 0 & 0 & \ldots & \ldots & 0 \\ p_1^{(1)} & p_0^{(1)} & 0 & \ldots & \ldots & 0 \\ p_2^{(1)} & p_1^{(1)} & p_0^{(1)} & 0 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ q_K^{(1)} & q_{K-1}^{(1)} & q_{K-2}^{(1)} & \ldots & \ldots & q_0^{(1)} \end{pmatrix} \begin{pmatrix} P_0^{(1)} \\ P_1^{(1)} \\ P_2^{(1)} \\ \vdots \\ P_K^{(1)} \end{pmatrix}. \tag{30}$$

The relations between $\mathbf{P}^{(2)}$ and $\mathbf{P}$, the probabilities of $S(t)$ at consecutive ordering times kMa^{-1}, $(k+1)Ma^{-1}$, are then given by

$$\mathbf{P}^{(2)} = \mathbf{p}^{(1)}\mathbf{p}\mathbf{P}, \tag{31}$$

where the matrix $\mathbf{p}^{(1)}\mathbf{p}$ is square. For ordering times, the stationary probability distribution $P_0, P_1, \ldots, P_K$ can be found by putting $P_i^{(2)} = P_i$ for $i = 0, 1, \ldots, K$, and solving the set of equations $\mathbf{P} = \mathbf{p}^{(1)}\mathbf{p}\mathbf{P}$ together with $\sum_{i=0}^{K} P_i = 1$. This solution may not be obtained easily in an explicit form, but it can always be computed numerically for any given distribution $\{p_i(\theta)\}$, $i = 0, 1, \ldots$. We note, in passing, that this stationary distribution for ordering times will be easier to compute if we first obtain the stationary probability distribution $P_0^{(1)}, P_1^{(1)}, \ldots, P_K^{(1)}$, for delivery times. Writing $P_i^{(3)} = \Pr\{S((k+1)Ma^{-1} + T) = i\}$ for the probability that $S(t)$ have a value i at delivery time $(k+1)Ma^{-1} + T$, we have that the relation between $\mathbf{P}^{(3)}$ and $\mathbf{P}^{(1)}$, the probabilities of $S(t)$ at consecutive delivery times $kMa^{-1} + T$, $(k+1)Ma^{-1} + T$, is given by

$$\mathbf{P}^{(3)} = \mathbf{p}\mathbf{p}^{(1)}\mathbf{P}^{(1)}, \tag{32}$$

where the matrix $\mathbf{p}\mathbf{p}^{(1)}$ is triangular. The stationary probability distribution for delivery times, $P_0^{(1)}, P_1^{(1)}, \ldots, P_K^{(1)}$, can be found by putting $P_i^{(3)} = P_i^{(1)}$ for $i = 0, 1, \ldots, K$, and solving the equations $\mathbf{P}^{(1)} = \mathbf{p}\mathbf{p}^{(1)}\mathbf{P}^{(1)}$, together with $\sum_{i=0}^{K} P_i^{(1)} = 1$. For numerical computation of $\mathbf{P}^{(1)}$, the work is considerably lightened by the fact that $\mathbf{p}\mathbf{p}^{(1)}$ is triangular; once the stationary probability distribution for delivery times is obtained, the simple relation $\mathbf{P} = \mathbf{p}^{(1)}\mathbf{P}^{(1)}$ will enable the stationary distribution for ordering times, $P_0, P_1, \ldots, P_K$ to be found.

(*b*) *A numerical example*

As an illustration of the computations necessary to evaluate the stationary distributions $\{P_i\}$ and $\{P_i^{(1)}\}$, for ordering and delivery times, we construct an example in which the consumption function $c(t)$ has, associated with it, a Poisson distribution such that

$$p_i(\theta) = e^{-a\theta}\frac{(a\theta)^i}{i!} \quad (i = 0, 1, \ldots).$$

We choose the maximum deficit $K = 9$, the time-lag between order and delivery $T = \frac{1}{2}$, the time interval between orders $Ma^{-1} = 1$, and the mean consumption per unit time $a = 4$. Then, since $T = Ma^{-1} - T = \frac{1}{2}$, we have that the p_i and $p_i^{(1)}$ defined previously are equal,

$$p_i = p_i^{(1)} = e^{-2}\,2^i/i! \quad (i = 0, 1, \ldots).$$

From Molina's tables for *Poisson's Exponential Binomial Limit*, we obtain

$$\left.\begin{aligned}
p_0 = p_0^{(1)} &= 0{\cdot}135335, \quad q_0 = q_0^{(1)} = 1{\cdot}000000,\\
p_1 = p_1^{(1)} &= 0{\cdot}270671, \quad q_1 = q_1^{(1)} = 0{\cdot}864665,\\
p_2 = p_2^{(1)} &= 0{\cdot}270671, \quad q_2 = q_2^{(1)} = 0{\cdot}593994,\\
p_3 = p_3^{(1)} &= 0{\cdot}180447, \quad q_3 = q_3^{(1)} = 0{\cdot}323324,\\
p_4 = p_4^{(1)} &= 0{\cdot}090224, \quad q_4 = q_4^{(1)} = 0{\cdot}142877,\\
p_5 = p_5^{(1)} &= 0{\cdot}036089, \quad q_5 = q_5^{(1)} = 0{\cdot}052653,\\
p_6 = p_6^{(1)} &= 0{\cdot}012030, \quad q_6 = q_6^{(1)} = 0{\cdot}016564,\\
p_7 = p_7^{(1)} &= 0{\cdot}003437, \quad q_7 = q_7^{(1)} = 0{\cdot}004534,\\
p_8 = p_8^{(1)} &= 0{\cdot}000859, \quad q_8 = q_8^{(1)} = 0{\cdot}001097,\\
&\qquad\quad q_9 = q_9^{(1)} = 0{\cdot}000237.
\end{aligned}\right\} \tag{33}$$

The equations giving the stationary distribution $\{P_i^{(1)}\}$ for delivery times are $\mathbf{P}^{(1)} = \mathbf{pp}^{(1)}\mathbf{P}$, together with $\sum_{i=0}^{K} P_i^{(1)} = 1$; these, in the case of the $\{p_i\}$ and $\{q_i\}$ given in (33), result in eleven equations, not all independent, for ten unknowns:

$$\begin{aligned}
&-0{\cdot}864460P_0^{(1)} + 0{\cdot}136284P_1^{(1)} + 0{\cdot}139256P_2^{(1)} + 0{\cdot}149657P_3^{(1)} + 0{\cdot}180862P_4^{(1)}\\
&\qquad + 0{\cdot}258876P_5^{(1)} + 0{\cdot}414902P_6^{(1)} + 0{\cdot}645941P_7^{(1)} + 0{\cdot}882981P_8^{(1)} + P_9^{(1)} = 0,\\
&0{\cdot}271117P_0^{(1)} - 0{\cdot}727584P_1^{(1)} + 0{\cdot}276590P_2^{(1)} + 0{\cdot}287625P_3^{(1)} + 0{\cdot}310012P_4^{(1)}\\
&\qquad + 0{\cdot}339183P_5^{(1)} + 0{\cdot}343934P_6^{(1)} + 0{\cdot}270671P_7^{(1)} + 0{\cdot}117019P_8^{(1)} = 0,\\
&0{\cdot}271486P_0^{(1)} + 0{\cdot}273334P_1^{(1)} - 0{\cdot}722144P_2^{(1)} + 0{\cdot}285591P_3^{(1)} + 0{\cdot}290341P_4^{(1)}\\
&\qquad + 0{\cdot}270671P_5^{(1)} + 0{\cdot}197408P_6^{(1)} + 0{\cdot}080388P_7^{(1)} = 0,\\
&0{\cdot}181348P_0^{(1)} + 0{\cdot}182615P_1^{(1)} + 0{\cdot}183837P_2^{(1)} - 0{\cdot}819553P_3^{(1)} + 0{\cdot}160777P_4^{(1)}\\
&\qquad + 0{\cdot}111934P_5^{(1)} + 0{\cdot}043756P_6^{(1)} = 0,\\
&0{\cdot}090630P_0^{(1)} + 0{\cdot}090224P_1^{(1)} + 0{\cdot}086834P_2^{(1)} + 0{\cdot}075304P_3^{(1)} - 0{\cdot}949117P_4^{(1)}\\
&\qquad + 0{\cdot}019336P_5^{(1)} = 0,\\
&0{\cdot}035683P_0^{(1)} + 0{\cdot}033922P_1^{(1)} + 0{\cdot}028904P_2^{(1)} + 0{\cdot}019136P_3^{(1)} + 0{\cdot}007126P_4^{(1)}\\
&\qquad - P_5^{(1)} = 0,\\
&0{\cdot}011129P_0^{(1)} + 0{\cdot}009368P_1^{(1)} + 0{\cdot}006111P_2^{(1)} + 0{\cdot}002242P_3^{(1)}\\
&\qquad - P_6^{(1)} = 0,\\
&0{\cdot}002623P_0^{(1)} + 0{\cdot}001692P_1^{(1)} + 0{\cdot}000614P_2^{(1)}\\
&\qquad - P_7^{(1)} = 0,\\
&0{\cdot}000413P_0^{(1)} + 0{\cdot}000148P_1^{(1)}\\
&\qquad - P_8^{(1)} = 0,\\
&0{\cdot}000032P_0^{(1)}\\
&\qquad - P_9^{(1)} = 0,\\
&P_0^{(1)} + P_1^{(1)} + P_2^{(1)} + P_3^{(1)} + P_4^{(1)}\\
&\qquad + P_5^{(1)} + P_6^{(1)} + P_7^{(1)} + P_8^{(1)} + P_9^{(1)} = 1.
\end{aligned}$$

Solving these equations by straightforward elimination, which is greatly simplified by the fact that $\mathbf{pp}^{(1)}$ is triangular, we obtain the following values for $\{P_i^{(1)}\}$:

$$\left.\begin{aligned} P_0^{(1)} &= 0{\cdot}148685, & P_5^{(1)} &= 0{\cdot}026833,\\ P_1^{(1)} &= 0{\cdot}281325, & P_6^{(1)} &= 0{\cdot}006382,\\ P_2^{(1)} &= 0{\cdot}277080, & P_7^{(1)} &= 0{\cdot}001036,\\ P_3^{(1)} &= 0{\cdot}177621, & P_8^{(1)} &= 0{\cdot}000103,\\ P_4^{(1)} &= 0{\cdot}080930, & P_9^{(1)} &= 0{\cdot}000005. \end{aligned}\right\} \tag{34}$$

From these we can compute the stationary distribution $\{P_i\}$ for ordering times, using the relations between $\mathbf{P}$ and $\mathbf{P}^{(1)}$ which are given in matrix form by $\mathbf{P} = \mathbf{p}^{(1)}\mathbf{P}^{(1)}$. This is easier than solving directly the equations $\mathbf{P} = \mathbf{p}^{(1)}\mathbf{pP}$ and $\sum_{i=0}^{K} P_i = 1$, where the matrix $\mathbf{p}^{(1)}\mathbf{p}$ is square. We obtain for $\{P_i\}$ the values:

$$\begin{aligned} P_0 &= 0{\cdot}020122, & P_5 &= 0{\cdot}154360,\\ P_1 &= 0{\cdot}078318, & P_6 &= 0{\cdot}099024,\\ P_2 &= 0{\cdot}153890, & P_7 &= 0{\cdot}053655,\\ P_3 &= 0{\cdot}202012, & P_8 &= 0{\cdot}025004,\\ P_4 &= 0{\cdot}198206, & P_9 &= 0{\cdot}015410. \end{aligned}$$

(c) *Comparison with Pitt's results for an infinite deficit*

It is interesting to compare the stationary distribution of the deficit $S(t)$ in the present case, where it is finite and restricted to the range $0 \leqslant S(t) \leqslant K$, with that obtained by Pitt for an infinite deficit with the same types of consumption and replacement functions. Pitt's problem for the infinite deficit, when the consumption function is of Poisson type, and the replacement function is the $r_2(t)$ defined in (4), is the exact analogue of the problem of the infinite dam considered in § IV (2). With a slight change in the notation of (22) to avoid confusion, we write the stationary distribution of the deficit $S(t)$ when its range is $0 \leqslant S(t) < \infty$ as $\{\Pi_i\}$ $(i = 0, 1, \ldots)$, where

$$\Pi_i = \sum_{v=0}^{i} p_{i-v} Q_v. \tag{35}$$

Here, the $p_{i-v} = e^{-aT}\dfrac{(aT)^{i-v}}{(i-v)!}$, and the Q_v are given by

$$\begin{aligned} Q_v &= M^{-1}\int_0^M e^{-u}\frac{u^v}{v!}du\\ &= M^{-1}\sum_{i=v+1}^{\infty} e^{-M}\frac{M^i}{i!}. \end{aligned}$$

We proceed to obtain a similar stationary distribution, also averaged over the time interval Ma^{-1}, for the present case of the finite deficit, and we shall write it $\{\Pi_i^{(1)}\}$ $(i = 0, 1, \ldots, K)$. To do this we consider the time interval θ, where $\theta \leqslant Ma^{-1}$ such that $kMa^{-1} + T + \theta$ is any time between two deliveries at times $kMa^{-1} + T$, and $(k+1)\,Ma^{-1} + T$. Let the probabilities that there be consumption i in this interval be

$$p_i(\theta) = e^{-a\theta}\frac{(a\theta)^i}{i!},$$

then if $\{P_i(kMa^{-1}+T+\theta)\}$ represent the distribution of the deficit $S(t)$ at the time $kMa^{-1}+T+\theta$, this will be given in matrix form by

$$\begin{pmatrix} P_0(kMa^{-1}+T+\theta) \\ P_1(kMa^{-1}+T+\theta) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ P_{K-1}(kMa^{-1}+T+\theta) \\ P_K(kMa^{-1}+T+\theta) \end{pmatrix} = \begin{pmatrix} p_0(\theta) & 0 & \cdots & \cdots & \cdots & 0 \\ p_1(\theta) & p_0(\theta) & 0 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{K-1}(\theta) & p_{K-2}(\theta) & \cdots & \cdots & p_0(\theta) & 0 \\ q_K(\theta) & q_{K-1}(\theta) & \cdots & \cdots & q_1(\theta) & q_0(\theta) \end{pmatrix} \begin{pmatrix} P_0^{(1)} \\ P_1^{(1)} \\ \vdots \\ P_{K-1}^{(1)} \\ P_K^{(1)} \end{pmatrix},$$

where $q_j(\theta) = \sum_{i=j}^{\infty} p_i(\theta)$, and $P_0^{(1)}, P_1^{(1)}, \ldots, P_K^{(1)}$, are the stationary probability distribution of the deficit at delivery times. We can write these in the forms

$$P_i(kMa^{-1}+T+\theta) = \sum_{v=0}^{i} P_{i-v}^{(1)} p_v(\theta) \quad (i = 0, 1, \ldots, K-1),$$

$$P_K(kMa^{-1}+T+\theta) = \sum_{v=0}^{K} P_{K-v}^{(1)} q_v(\theta).$$

The stationary distribution $\{\Pi_i^{(1)}\}$ will then be obtained from the $\{P_i(kMa^{-1}+T+\theta)\}$ by averaging them over the interval of time Ma^{-1}, so that

$$\left.\begin{aligned} \Pi_i^{(1)} &= aM^{-1}\int_0^{Ma^{-1}} \sum_{v=0}^{i} P_{i-v}^{(1)} p_v(\theta)\, d\theta, \\ &= \sum_{v=0}^{i} P_{i-v}^{(1)} Q_v \quad (i = 0, 1, \ldots, K-1); \\ \Pi_K^{(1)} &= aM^{-1}\int_0^{Ma^{-1}} \sum_{v=0}^{K} P_{K-v}^{(1)} q_v(\theta)\, d\theta, \\ &= \sum_{v=0}^{K} P_{K-v}^{(1)} \sum_{j=v}^{\infty} Q_j, \end{aligned}\right\} \tag{36}$$

where there is some similarity of form with (35).

In the case of our numerical example (§ V (2)(*b*)), where $K = 9$, $T = \frac{1}{2}$, $M = a = 4$, the values of the p_{i-v} in (35) are given by equation (33) for values of $i = 0, 1, \ldots, 8$; the values of the Q_v in (35) and (36) are given for $v = 0, 1, \ldots, 8$ by

$$\begin{aligned} Q_0 &= 0{\cdot}245421, \quad Q_3 = 0{\cdot}141633, \quad Q_6 = 0{\cdot}027669, \\ Q_1 &= 0{\cdot}227106, \quad Q_4 = 0{\cdot}092791, \quad Q_7 = 0{\cdot}012784, \\ Q_2 &= 0{\cdot}190474, \quad Q_5 = 0{\cdot}053718, \quad Q_8 = 0{\cdot}005341; \end{aligned}$$

and the values of the $\{P_i^{(1)}\}$ are given by (34). Using these, we compute

$$\begin{aligned} \Pi_0 &= 0{\cdot}033214, \quad \Pi_0^{(1)} = 0{\cdot}036490, \\ \Pi_1 &= 0{\cdot}097164, \quad \Pi_1^{(1)} = 0{\cdot}102810, \\ \Pi_2 &= 0{\cdot}153677, \quad \Pi_2^{(1)} = 0{\cdot}160212, \\ \Pi_3 &= 0{\cdot}176480, \quad \Pi_3^{(1)} = 0{\cdot}181162, \\ \Pi_4 &= 0{\cdot}165573, \quad \Pi_4^{(1)} = 0{\cdot}166619, \\ \Pi_5 &= 0{\cdot}134440, \quad \Pi_5^{(1)} = 0{\cdot}132132, \\ \Pi_6 &= 0{\cdot}097291, \quad \Pi_6^{(1)} = 0{\cdot}093169, \\ \Pi_7 &= 0{\cdot}063731, \quad \Pi_7^{(1)} = 0{\cdot}059328, \\ \Pi_8 &= 0{\cdot}038132, \quad \Pi_8^{(1)} = 0{\cdot}034385, \\ \sum_{i=9}^{\infty} \Pi_i &= 0{\cdot}040298, \quad \Pi_9^{(1)} = 0{\cdot}033693. \end{aligned}$$

There would appear to be a close similarity in the values of the $\{\Pi_i\}$ and $\{\Pi_i^{(1)}\}$, so that we could use the more easily computed values of Π_i as approximations for the $\Pi_i^{(1)}$; however, it is not possible without further work to state within what bounds this approximation would be valid.

(3) *Provisioning with orders at fixed times $t = kMa^{-1}$, and replacements at fixed times $t = (k+1)Ma^{-1}$ $(k = 1, 2, \ldots)$, with time-lag Ma^{-1}*

(*a*) *General equations, and an exact solution for the case of consumption with a discrete geometric distribution*

This is a special case of the previous problem (§ V (2) (*a*)), when the times for delivery of the replacements and for ordering coincide; that is, a new order for replacements is put in every time a delivery occurs, at time intervals Ma^{-1}. The consumption function is of the same form (25), the replacement function is $r_2(t)$, (4), the overflow function is defined by (26), where the time-lag T is Ma^{-1}; in these conditions, with the same notation as before, the equations (28) relate the probability distributions $\{P_i\}$ and $\{P_i^{(1)}\}$ for $S(t)$ at consecutive ordering and replacement times kMa^{-1}, and $(k+1)Ma^{-1}$ $(k = 1, 2, \ldots)$. To obtain the stationary probability distribution $P_0, P_1, \ldots, P_K$, for these times, we write $P_i^{(1)} = P_i$ $(i = 0, 1, \ldots, K)$, and solve the equations $\mathbf{P} = \mathbf{pP}$ together with $\sum_{i=0}^{K} P_i = 1$.

If, for convenience, we group these equations in a slightly different way in the top half, we have

$$\left.\begin{aligned} P_0 &= p_0(P_0 + \ldots + P_K) + q_1 P_K, \\ P_1 &= p_1(P_0 + \ldots + P_{K-1}) + q_2 P_{K-1}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ P_{K-1} &= p_{K-1}P_0 + q_{K-1}P_1, \\ P_K &= q_K P_0, \\ 1 &= P_0 + P_1 + \ldots + P_K, \end{aligned}\right\} \tag{37}$$

which can be solved in pairs, starting with P_0 and P_K, and working towards the centre of the group in the order P_1, P_{K-1}; P_2, P_{K-2}; etc. We obtain the recurrence relations

$$\begin{aligned} P_0 &= p_0(1 - q_1 q_K)^{-1}, \\ P_K &= q_K P_0 = q_K p_0 (1 - q_1 q_K)^{-1}; \\ P_1 &= \{p_1(1 - P_K) + q_2 p_{K-1} P_0\}(1 - q_2 q_{K-1})^{-1}, \\ P_{K-1} &= p_{K-1}P_0 + q_{K-1}P_1; \end{aligned}$$

for the first two pairs of solutions. Assuming that all probabilities P_n from $P_0, P_1, \ldots$ to P_{i-1}, and $P_K, P_{K-1}, \ldots$ to P_{K-i+1} are known, the following formulae enable all other P_i to be found from these by a repetition of the process:

$$\left.\begin{aligned} P_i &= \left\{p_i\left(1 - \sum_{n=K-i+1}^{K} P_n\right) + q_{i+1} p_{K-i} \sum_{n=0}^{i-1} P_n\right\}(1 - q_{i+1} q_{K-i})^{-1}, \\ P_{K-i} &= p_{K-i} \sum_{n=0}^{i-1} P_n + q_{K-i} P_i \quad (i = 1, \ldots, K-1). \end{aligned}\right\} \tag{38}$$

A distribution $\{p_i\}$ which gives a simple explicit solution of these equations is the geometric; it is important to note, however, that this does not satisfy the independence condition for non-overlapping intervals of time postulated in (25). For the geometric distribution

$$p_i = AB^i \quad (i = 0, 1, \ldots), \tag{39}$$

where $A+B=1$, we find from the recurrence relations (38) that the stationary distribution of the deficit $\{P_i\}$ is given by

$$P_i = AB^i(1-B^{K+1})^{-1} \quad (i = 0, 1, \ldots, K). \tag{40}$$

(*b*) *An exact solution for the case of consumption with a continuous exponential distribution*

Suppose that instead of dealing with integral units of the deficit, consumption and replacement, we deal with units of magnitude Δ; Δ will tend to zero, and K and i will tend to infinity in such a way that the maximum deficit $K\Delta$ tends to the value k, and a deficit, consumption or replacement $i\Delta$ tends to the value s $(i = 0, 1, \ldots, K)$. The distribution of the consumption then becomes continuous; for if we put

$$A = 1-e^{-\mu\Delta},$$
$$B = e^{-\mu\Delta},$$

and let $\Delta \to 0$ in (39), so that $i\Delta \to s$, we have that the distribution of the consumption in the interval of time Ma^{-1} is

$$\begin{aligned} p(s)\,ds &= \lim_{\Delta\to 0} \Delta^{-1} p_i\,ds \\ &= \lim_{\Delta\to 0} \Delta^{-1} e^{-\mu i\Delta}(1-e^{-\mu\Delta})\,ds \\ &= \mu e^{-\mu s}\,ds, \end{aligned} \tag{41}$$

which is continuous, and of the exponential form. The solutions (40) of equations (37) will lead by the same limiting process to an exact solution of the following provisioning problem.

A deficit $S(t)$ is defined continuously in the range $0 \leqslant S(t) \leqslant k$, at the fixed times nMa^{-1} $(n = 0, 1, \ldots)$, and is subject to the following conditions:

(1) a consumption function $c(t)$ such that in a time interval Ma^{-1} the distribution of the consumption is continuous and of the exponential type

$$\begin{aligned} p(s)\,ds &= \Pr\{s < c((n+1)\,Ma^{-1}) - c(nMa^{-1}) < s+ds\} \\ &= \mu e^{-\mu s}\,ds; \end{aligned}$$

(2) a replacement function such that

$$r_2((n+1)\,Ma^{-1}) = c(nMa^{-1}),$$

where irregular quantities lying between the values $(s, s+ds)$ equal to the consumption in the previous time interval Ma^{-1} are ordered at times nMa^{-1} and delivered and added to the stock at times $(n+1)\,Ma^{-1}$;

(3) an overflow function such that in any time interval Ma^{-1} between ordering and delivery times nMa^{-1}, $(n+1)\,Ma^{-1}$, the consumption is lost at any time after the deficit rises to the value k, and is otherwise met from the stock.

For small Δ, the equations (37) hold, with the proviso that $\{p_i\}$ represent the probabilities that the consumption in an interval Ma^{-1} is $i\Delta$, and $\{P_i\}$ the stationary probabilities that the deficit have a value $i\Delta$, so that the solutions (40) to them also hold. In the limit as Δ tends to zero in the manner defined above, the continuous stationary distribution of the deficit $S(t)$ is given by

$$\begin{aligned} f(s)\,ds &= \lim_{\Delta\to\infty} \Delta^{-1} P_i\,ds \\ &= \mu e^{-\mu s}(1-e^{-\mu k})^{-1}\,ds \quad (0 \leqslant s \leqslant k), \end{aligned} \tag{42}$$

a truncated exponential distribution.

This solution can also be obtained by taking the limit as Δ tends to zero for equations (37); this gives an integral equation for the continuous stationary probability distribution of the deficit,

$$f(s)\,ds = p(s)\,ds\int_0^{k-s} f(u)\,du + f(k-s)\,ds\int_s^{\infty} p(u)\,du. \tag{43}$$

It is easily verified that (42) will satisfy this equation.

VI. An exact solution of the storage problem with a Poisson input

In attempting to find algebraic solutions for the stationary probability distribution $\{P_i\}$ of the dam content $Z(t)$ in Moran's problem (§ II), where $\mathbf{P} = \mathbf{pP}$, $\mathbf{P}$ being the column vector of the P_i, it is found that no simple solution in an explicit form exists for an input function of the Poisson type with parameter a, such that in the interval of time t, $t+1$, the probabilities of an input $X(t)$ being i units is

$$p_i = e^{-a}a^i/i! \quad (i = 0, 1, \ldots).$$

However, if in this problem the conditions for the input of Poisson type and for the overflow are left unchanged, but the release rule is altered so that instead of the release rule (3) of § II (M units at times $t+1$ if the total content of the dam $Z(t)+X(t)$ is greater or equal to M, or the total content of the dam $Z(t)+X(t)$ if this is less than M), a new release rule is prescribed so that there is an output with a steady flow when the dam contains water, and no output when it is empty, then $Z(t)$ is continuous, and it is possible to obtain solutions in an explicit form for its stationary distribution.

Our method of approach will be to consider Moran's set of equations (5) for an infinitesimal interval of time δt, instead of a unit interval of time; the input will remain of Poisson type, and be a multiple of a definite discrete unit which will not tend to zero, while a small discrete release will be made at the end of each infinitesimal interval of time under a rule of the same type as rule (3) of § II. The equations can then be solved explicitly, and when $\delta t \to 0$, these solutions will provide the continuous stationary probability distribution for the dam content $Z(t)$, which is then itself continuous, when the release rule prescribed gives an output with a steady flow when the dam contains water, or no output when the dam is empty.

Before proceeding further, it is preferable, since the theory of dams was identified with the general storage theory, to view the problem as one of the general storage function. It then becomes possible to interpret the problem as one of provisioning as well; here we would have a consumption of Poisson type, a replacement rule such that there is continuous delivery at a steady rate when the stock is not at its maximum value, and zero when it is, and an overflow function such that the consumption is lost when there is no stock to meet it. In practice, this could conceivably be the case of a stock of grain in a silo, where the replacement is a steady flow of grain which is stopped when the silo is full, and the consumption of Poisson type is lost when the silo is empty, and met from the stock otherwise.

We frame the conditions of the problem for the storage function as follows: a discrete storage function $S(t)$ is defined in the range $0 \leqslant S(t) \leqslant (K-1)\Delta$, where $K = bH + U$ (b, H, U integers) at the fixed times $t = 0, \delta t, 2\delta t, \ldots$, and is subject to the conditions that

(1) it is fed by a discrete input function $I(t)$ of Poisson type with parameter a, such that in a small interval of time δt there may be inputs of $H\Delta$ or of no units respectively with probabilities

$$p = \Pr\{I(t+\delta t) - I(t) = H\Delta\} = 1 - e^{-a\delta t},$$
$$q = \Pr\{I(t+\delta t) - I(t) = 0\} \quad = e^{-a\delta t};$$

(2) the output function is such that at fixed times $(n+1)\,\delta t$ $(n = 0, 1, \ldots)$, the discrete output consists of a quantity $\Delta = C\,\delta t$ units, where C is a constant, provided that the storage function at a time just previous to $(n+1)\,\delta t$ is greater than Δ,

$$S(n\,\delta t) + I((n+1)\,\delta t) - I(n\,\delta t) \geqslant \Delta,$$

but is zero otherwise;

(3) the overflow function is such that in the interval $n\,\delta t$, $(n+1)\,\delta t$ there is no overflow if

$$S(n\,\delta t) + I((n+1)\,\delta t) - I(n\,\delta t) \leqslant K\Delta,$$

but there is an overflow

$$S(n\,\delta t) + I((n+1)\,\delta t) - I(n\,\delta t) - K\Delta,$$

if this is greater than zero.

Writing P_i, $P_i^{(1)}$ for the probabilities that $S(t)$ take the values $i\Delta$ at times $n\,\delta t$, $(n+1)\,\delta t$ respectively, where $i = 0, 1, \ldots, K-1$, we have that the relations between these probabilities are

$$\left.\begin{aligned}
P_0^{(1)} &= qP_0 + qP_1,\\
P_1^{(1)} &= qP_2,\\
&\cdots\cdots\cdots\cdots\\
P_{H-2}^{(1)} &= qP_{H-1},\\
P_{H-1}^{(1)} &= pP_0 + qP_H,\\
P_H^{(1)} &= pP_1 + qP_{H+1},\\
&\cdots\cdots\cdots\cdots\\
P_{2H-2}^{(1)} &= pP_{H-1} + qP_{2H-1},\\
&\cdots\cdots\cdots\cdots\\
P_{jH-1}^{(1)} &= pP_{(j-1)H} + qP_{jH},\\
&\cdots\cdots\cdots\cdots\\
P_{(j+1)H-2}^{(1)} &= pP_{jH-1} + qP_{(j+1)H-1},\\
&\cdots\cdots\cdots\cdots\\
P_{bH-1}^{(1)} &= pP_{(b-1)H} + qP_{bH},\\
&\cdots\cdots\cdots\cdots\\
P_{K-1}^{(1)} &= P_{bH+U-1}^{(1)} = p(P_{(b-1)H+U} + \ldots + P_{bH+U-1}),
\end{aligned}\right\} \quad (44)$$

or in matrix form, $\mathbf{P}^{(1)} = \mathbf{pP}$, where $\mathbf{P}$, $\mathbf{P}^{(1)}$ are column vectors with elements P_i, $P_i^{(1)}$ respectively, and $\mathbf{p}$ is the matrix of coefficients.

Formally, this is Moran's set of equations (5), in which we have set $p_0 = q$, $p_H = p$, and all other $p_i = 0$ for i not equal to 0 or H. For the stationary distribution of $S(t)$, we put $P_i^{(1)} = P_i$, and solve the equations $\mathbf{P} = \mathbf{pP}$, together with $\sum_{i=0}^{K-1} P_i = 1$. We see clearly that the solution of these equations will fall into the $(b+1)$ distinct classes:

(0)th class: $P_1, P_2, \ldots, P_{H-1}$;

(1)th class: $P_H, P_{H+1}, \ldots, P_{2H-1}$;

$\cdots\cdots\cdots\cdots$

(j)th class: $P_{jH}, P_{jH+1}, \ldots, P_{(j+1)H-1}$;

$\cdots\cdots\cdots\cdots$

(b)th class: $P_{bH}, P_{bH+1}, \ldots, P_{K-1}$.

The solution for the 0th class is immediately obvious from (44),

$$P_R = pP_0q^{-R} \quad (R = 1, 2, \ldots, H-1); \tag{45}$$

for the jth class, we obtain from our equations (44) the equation

$$P_{jH+R} = q^{-(R+1)}\left\{P_{jH-1} - p\sum_{N=0}^{R} q^N P_{(j-1)H+N}\right\} \quad (j = 1, \ldots, b;\ R = 0, \ldots, H-1), \tag{46}$$

relating any probability P_{jH+R} in the jth class with the probabilities $P_{(j-1)H+N}$, P_{jH-1} of the $(j-1)$th class, so that starting with the solutions (45) for the 0th class, we can obtain solutions progressively for one class after another.

Consider the formula

$$P_{jH+R} = pP_0q^{-(jH+R)}\left(\sum_{v=0}^{j}(-1)^v p^{v-1}q^{v(H-1)}\left\{\binom{(j-v)H+R+v-1}{v-1} + p\binom{(j-v)H+R+v-1}{v}\right\}\right), \tag{47}$$

where $R = 1, 2, \ldots, H-1$ for $j = 0$, and $R = 0, 1, \ldots, H-1$ for $j = 1, 2, \ldots, b$, and where it is understood that $\binom{n}{r} = 0$ if $r < 0$, or if $n < r$. For $j = 0$, this gives

$$P_R = pP_0q^{-R} \quad (R = 1, 2, \ldots, H-1),$$

which is the solution (45) for the (0)th class. We prove that formula (47) holds for all values of j by induction; assume that it holds for the class $(j-1)$, then substituting for $P_{(j-1)H+N}$ and P_{jH-1} given by (47) in the equation (46), we obtain

$$\begin{aligned}
P_{jH+R} = {} & q^{-(R+1)}\left\{pP_0q^{-(jH-1)}\left(\sum_{v=0}^{j-1}(-1)^v p^{v-1}q^{v(H-1)}\left\{\binom{(j-v)H+v-2}{v-1} + p\binom{(j-v)H+v-2}{v}\right\}\right)\right. \\
& - p^2P_0q^{-(j-1)H}\sum_{N=0}^{R}\left(\sum_{v=0}^{j-1}(-1)^v p^{v-1}q^{v(H-1)}\right. \\
& \left.\left.\times\left\{\binom{(j-1-v)H+N+v-1}{v-1} + p\binom{(j-1-v)H+N+v-1}{v}\right\}\right)\right\} \\
= {} & pP_0q^{-(jH+R)}\left\{1 + \sum_{v=1}^{j-1}(-1)^v p^{v-1}q^{v(H-1)}\left(\left\{\binom{(j-v)H+v-2}{v-1}\right.\right.\right. \\
& \left.\left.\left. + \sum_{N=0}^{R}\binom{(j-v)H+N+v-2}{v-2}\right\} + p\left\{\binom{(j-v)H+v-2}{v} + \sum_{N=0}^{R}\binom{(j-v)H+N+v-2}{v-1}\right\}\right)\right. \\
& \left. + (-1)^j p^{j-1}q^{j(H-1)}\left\{\sum_{N=0}^{R}\binom{N+j-2}{j-2} + p\sum_{N=0}^{R}\binom{N+j-2}{j-1}\right\}\right\}.
\end{aligned} \tag{48}$$

Now in (48), consider

$$\left\{\binom{(j-v)H+v-2}{v-1} + \sum_{N=0}^{R}\binom{(j-v)H+N+v-2}{v-2}\right\};$$

since

$$\sum_{t=0}^{n}\binom{t}{w} = \sum_{t=w}^{n}\binom{t}{w} = \binom{n+1}{w+1},$$

then

$$\begin{aligned}
& \left\{\binom{(j-v)H+v-2}{v-1} + \sum_{N=0}^{R}\binom{(j-v)H+N+v-2}{v-2}\right\} \\
& = \left\{\binom{(j-v)H+v-2}{v-1} + \sum_{N=-(j-v)H}^{R}\binom{(j-v)H+N+v-2}{v-2} - \sum_{N=-(j-v)H}^{-1}\binom{(j-v)H+N+v-2}{v-2}\right\} \\
& = \binom{(j-v)H+R+v-1}{v-1}.
\end{aligned}$$

In exactly the same way, we reduce the other terms in (48), so that

$$\left\{\binom{(j-v)H+v-2}{v}+\sum_{N=0}^{R}\binom{(j-v)H+N+v-2}{v-1}\right\}=\binom{(j-v)H+R+v-1}{v},$$

and

$$\sum_{N=0}^{R}\binom{N+j-2}{j-2}+p\sum_{N=0}^{R}\binom{N+j-2}{j-1}=\binom{R+j-1}{j-1}+p\binom{R+j-1}{j}.$$

We thus obtain for P_{jH+R} in (48) the following expression:

$$P_{jH+R}=pP_0q^{-(jH+R)}\times\left\{1+\sum_{v=1}^{j}(-1)^v p^{v-1}q^{v(H-1)}\left(\binom{(j-v)H+R+v-1}{v-1}+p\binom{(j-v)H+R+v-1}{v}\right)\right\}, \quad (49)$$

which is equivalent to (47). Now, since this has been shown to hold for the 0th class, then it will hold for the 1th class, and so on up to class (b); this proves that (47) (or the slightly different form (49)) is the solution required for the stationary distribution of $S(t)$ in the discrete case considered. We write out in full the formulae for a few classes so as to give an idea of their structure:

$$P_R = pP_0q^{-R} \quad (R=1,2,\ldots,H-1);$$

$$P_{H+R} = pP_0q^{-(H+R)}\left\{1-q^{H-1}\left(1+p\binom{R}{1}\right)\right\},$$

$$P_{2H+R} = pP_0q^{-(2H+R)}\left\{1-q^{H-1}\left(1+p\binom{H+R}{1}\right)+pq^{2(H-1)}\left(\binom{R+1}{1}+p\binom{R+1}{2}\right)\right\},$$

$$P_{3H+R} = pP_0q^{-(3H+R)}\left\{1-q^{H-1}\left(1+p\binom{2H+R}{1}\right)+pq^{2(H-1)}\left(\binom{H+R+1}{1}+p\binom{H+R+1}{2}\right)\right.$$
$$\left.-p^2q^{3(H-1)}\left(\binom{R+2}{2}+p\binom{R+2}{3}\right)\right\} \quad (R=0,1,\ldots,H-1).$$

For the complete solution in any particular case, the value of P_0 is also required; this can be found from the equation $\sum_{i=0}^{K-1}P_i=1$.

We now proceed to the continuous case by letting $\delta t\to 0$, and consequently $\Delta=C\,\delta t\to 0$, where the constant C can now be interpreted as the rate of output per unit time. We allow Δ to tend to zero, and the previously defined integers K, H and R to tend to infinity in such a way that

(1) the maximum value of the storage function $K\Delta\to k$, or $(bH+U)\Delta\to bh+u$,

(2) the input $H\Delta\to h$,

(3) the values of the storage function $(jH+R)\Delta\to jh+r$, which we shall also write s $(j=0,1,\ldots,b;\ 0\leqslant r<h)$.

As a result of this limiting process, the conditions of the problem which we have just considered for the storage function are changed; the function $S(t)$ has now a continuous range of values, $0\leqslant S(t)<k$, where $k=bh+u$, and is defined for all time t (t continuous) in such a way that

(1) it is fed by a discrete input of Poisson type with parameter a, so that in a small interval of time δt there may be inputs of h or of no units with probabilities

$$p=\Pr\{I(t+\delta t)-I(t)=h\}=1-e^{-a\delta t}=1-e^{-(a/C)\Delta}=1-e^{-\mu\Delta},$$
$$q=\Pr\{I(t+\delta t)-I(t)=0\}=e^{-a\delta t}\quad=e^{-(a/C)\Delta}\quad=e^{-\mu\Delta},$$

where $\mu=a/C$.

(2) the output is such that it has a constant rate C per unit time, provided the storage function is not zero, but is zero if the storage function is zero;

(3) the overflow is such that in any interval of time δt there is no overflow if

$$S(t)+I(t+\delta t)-I(t)\leqslant k,$$

but there is an overflow

$$S(t)+I(t+\delta t)-I(t)-k$$

if this is greater than zero.

We write the continuous stationary probability density for $S(t)$ as $f(s)$, so that

$$f(s)\,ds=\Pr\{s<S(t)<s+ds\}.$$

Now for $s=\lim_{\Delta\to 0}(jH+R)\Delta=jh+r$, we have that

$$f(s)\,ds=\lim_{\Delta\to 0}P_{jH+R}\Delta^{-1}ds;$$

this enables us to obtain from (45) and (49) the equations for $f(s)$ in various ranges of magnitude h. For $j=0$ $(0<r<h)$ we have

$$\begin{aligned}f(r)&=\lim_{\Delta\to 0}\Delta^{-1}pP_0q^{-R}=\lim_{\Delta\to 0}P_0\Delta^{-1}(1-e^{-\mu\Delta})\,e^{\mu\Delta R}\\&=\mu\,e^{\mu r}P_0.\end{aligned}\tag{50}$$

For $j=1,2,\ldots,b$ and $0\leqslant r<h$, we have

$$\begin{aligned}f(jh+r)&=\lim_{\Delta\to 0}\Delta^{-1}pP_0q^{-(jH+R)}\Bigg(1+\sum_{v=1}^{j}(-1)^v p^{v-1}q^{v(H-1)}\\&\qquad\times\left\{\binom{(j-v)H+R+v-1}{v-1}+p\binom{(j-v)H+R+v-1}{v}\right\}\Bigg)\\&=\lim_{\Delta\to 0}\Delta^{-1}(1-e^{-\mu\Delta})\,e^{\mu\Delta(jH+R)}P_0\Bigg(1+\sum_{v=1}^{j}(-1)^v e^{-\mu v\Delta(H-1)}\\&\qquad\times\left\{(\mu\Delta)^{v-1}\binom{(j-v)H+R+v-1}{v-1}+(\mu\Delta)^v\binom{(j-v)H+R+v-1}{v}\right\}\Bigg)\\&=\mu\,e^{\mu(jh+r)}P_0\Bigg(1+\sum_{v=1}^{j}(-1)^v e^{-\mu vh}\left\{\mu^{v-1}\frac{((j-v)h+r)^{v-1}}{(v-1)!}+\frac{\mu^v((j-v)h+r)^v}{v!}\right\}\Bigg).\end{aligned}\tag{51}$$

In addition to these probability densities, which are continuous in the intervals in which they are defined, there is also a concentration of probability P_0 at $s=0$. To complete the solution, we require the value of P_0; this will be given in any particular case where the value of $k=bh+u$ is known by

$$P_0+\sum_{j=0}^{b-1}\int_{r=0}^{h}f(jh+r)\,dr+\int_{r=0}^{u}f(bh+r)\,dr=1.\tag{52}$$

We write out in full some of the formulae for the $f(jh+r)$, in order to indicate their structure:

$$\left.\begin{aligned}f(r)&=\mu\,e^{\mu r}P_0\quad(0<r<h),\\f(h+r)&=\mu\,e^{\mu(h+r)}P_0\{1-e^{-\mu h}(1+\mu r)\},\\f(2h+r)&=\mu\,e^{\mu(2h+r)}P_0\left\{1-e^{-\mu h}(1+\mu(h+r))+\mu\,e^{-2\mu h}\left(r+\mu\frac{r^2}{2!}\right)\right\},\\f(3h+r)&=\mu\,e^{\mu(3h+r)}P_0\left\{1-e^{-\mu h}(1+\mu(2h+r))+\mu\,e^{-2\mu h}\left((h+r)+\mu\frac{(h+r)^2}{2!}\right)\right.\\&\qquad\left.-\mu^2e^{-3\mu h}\left(\frac{r^2}{2!}+\mu\frac{r^3}{3!}\right)\right\}\quad(0\leqslant r<h).\end{aligned}\right\}\tag{53}$$

It is interesting to note that these formulae for $f(jh+r)$ $(j = 1, 2, \ldots, b)$ can also be obtained from the limit of equation (46), when $\Delta \to 0$,

$$f(jh+r) = \lim_{\Delta \to 0} P_{jH+R}\Delta^{-1}$$
$$= \lim_{\Delta \to 0} q^{-(R+1)}\left\{P_{jH-1}\Delta^{-1} - p\Delta^{-1}\sum_{N=0}^{R} q^N (P_{(j-1)H+N}\Delta^{-1})\Delta\right\}. \quad (54)$$

Now for $j = 1$, we have

$$f(h+r) = \lim_{\Delta \to 0} P_{H+R}\Delta^{-1}$$
$$= \lim_{\Delta \to 0} q^{-(R+1)}\left\{P_{H-1}\Delta^{-1} - p\Delta^{-1}\sum_{N=1}^{R} q^N (P_N\Delta^{-1})\Delta - p\Delta^{-1}P_0\right\}$$
$$= e^{\mu r} f(h-0) - \mu e^{\mu r}\left\{\int_0^r e^{-\mu n} f(n)\, dn + P_0\right\}, \quad (55)$$

since $P_N\Delta^{-1} \to f(n)$, where $N\Delta \to n$ for all $N \neq 0$, but there is a concentration of probability P_0 at $s = 0$. For $j = 2, 3, \ldots, b$ the equation (54) gives

$$f(jh+r) = e^{\mu r} f(jh-0) - \mu e^{\mu r}\int_0^r e^{-\mu n} f((j-1)h+n)\, dn. \quad (56)$$

The formulae (53) can be obtained by putting the formula (50) for $f(r)$ in (55) and integrating it to find $f(h+r)$, then repeating this process for $f(jh+r)$ by using (56). More generally, it can be proved by induction that our solutions (50) and (51) satisfy the integral equations (55) and (56).

It is found that (52) does not lend itself to the computation of P_0 as well as an alternative equation for the integral of the probabilities, which we now obtain. To do so, we write the probability density $f(s)$ for the various intervals in the following manner:

$$f(s) = \mu e^{\mu s} P_0 \quad (0 < s < h);$$

$$f(s) = \mu e^{\mu s} P_0\left(1 + \sum_{v=1}^{j} (-1)^v e^{-\mu v h}\left\{\mu^{v-1}\frac{(s-vh)^{v-1}}{(v-1)!} + \mu^v \frac{(s-vh)^v}{v!}\right\}\right)$$
$$(jh \leqslant s < (j+1)h;\ j = 1, 2, \ldots, b).$$

The equation (52) can then be written as

$$P_0\left(1 + \int_0^{bh+u} \mu e^{\mu s}\, ds + \sum_{j=1}^{b} (-1)^j \int_{s=jh}^{bh+u} \mu e^{\mu(s-jh)}\left\{\mu^{j-1}\frac{(s-jh)^{j-1}}{(j-1)!} + \mu^j \frac{(s-jh)^j}{j!}\right\} ds\right) = 1,$$

which shortens considerably the work involved in the computation of P_0.

I am greatly indebted to Prof. P. A. P. Moran for suggesting several of the problems, and for his extremely helpful discussion and criticism of the work throughout all stages.

REFERENCES

Feller, W. (1950). *An Introduction to Probability Theory and its Applications.* New York: John Wiley and Sons.

Frazer, R. A., Duncan, W. J. & Collar, A. R. (1947). *Elementary Matrices.* Cambridge University Press.

Moran, P. A. P. (1954). A probability theory of dams and storage systems. *Aust. J. Appl. Sci.* **5**, no. 2, pp. 116–24.

Moran, P. A. P. (1955). A probability theory of dams and storage systems. II (to appear).

Pitt, H. R. (1946). A theorem on random functions with applications to a theory of provisioning. *J. Lond. Math. Soc.* **21**, 16–22.

APPROXIMATE CONFIDENCE INTERVALS

III. A BIAS CORRECTION

By M. S. BARTLETT
University of Manchester

In my second paper on approximate confidence intervals (Bartlett, 1953, to be referred to as II), where a general method of eliminating 'nuisance parameters' was discussed, it was claimed (see top of p. 310) that the effect of substituting maximum-likelihood estimates for the nuisance parameters could be neglected to the first order ($O(1/\sqrt{n})$) of approximation. I regret to say that this assertion is not in general correct, and the statistic T_e obtained after such substitution may have a bias of $O(1/\sqrt{n})$ which requires correction. It is even necessary in the expansion

$$T_e = T + (\hat{\theta}_2 - \theta_2)\frac{\partial T}{\partial \theta_2} + \tfrac{1}{2}(\hat{\theta}_2 - \theta_2)^2 \frac{\partial^2 T}{\partial \theta_2^2} + \dots \tag{1}$$

to make allowance not only for the second term, but also for the *third* term, on the right-hand side of (1), or to use the equivalent approximation to this order given in equation (14) of II. For convenience we may write (cf. equation (11) of II)

$$T \equiv a\frac{\partial L}{\partial \theta_1} + b\frac{\partial L}{\partial \theta_2}, \tag{2}$$

where $a \equiv 1/\sqrt{I_{11.2}}$, $b \equiv -I_{12}/(I_{22}\sqrt{I_{11.2}})$. Then, noting that $\partial T/\partial\theta_2$ is $O(1)$, and $\partial^2 T/\partial\theta_2^2$ $O(\sqrt{n})$, we find, omitting any terms of smaller order of magnitude than $1/\sqrt{n}$,

$$\begin{aligned} T_e - T = {} & \frac{1}{I_{22}}\frac{\partial L}{\partial \theta_2}\left[a\frac{\partial^2 L}{\partial\theta_1\partial\theta_2} + b\frac{\partial^2 L}{\partial\theta_2^2} + \frac{\partial L}{\partial\theta_1}\frac{\partial a}{\partial\theta_2} + \frac{\partial L}{\partial\theta_2}\frac{\partial b}{\partial\theta_2}\right] \\ & + \frac{1}{2I_{22}^2}\left(\frac{\partial L}{\partial\theta_2}\right)^2\left[aE\left(\frac{\partial^3 L}{\partial\theta_1\partial\theta_2^2}\right) + bE\left(\frac{\partial^3 L}{\partial\theta_2^3}\right) - 2I_{12}\frac{\partial a}{\partial\theta_2} - 2I_{22}\frac{\partial b}{\partial\theta_2}\right]. \end{aligned} \tag{3}$$

Hence, making use of relations such as

$$E\left(\frac{\partial L}{\partial\theta_2}\frac{\partial^2 L}{\partial\theta_1\partial\theta_2}\right) = -E\left(\frac{\partial^3 L}{\partial\theta_1\partial\theta_2^2}\right) - \frac{\partial I_{12}}{\partial\theta_2}, \tag{4}$$

we obtain from (3)

$$\begin{aligned} E\{T_e\} = {} & -\frac{a}{2I_{22}}\left[E\left(\frac{\partial^3 L}{\partial\theta_1\partial\theta_2^2}\right) + 2\frac{\partial I_{12}}{\partial\theta_2}\right] \\ & - \frac{b}{2I_{22}}\left[E\left(\frac{\partial^3 L}{\partial\theta_2^3}\right) + 2\frac{\partial I_{22}}{\partial\theta_2}\right] + O\left(\frac{1}{n}\right). \end{aligned} \tag{5}$$

When $I_{12} = 0$, this reduces to

$$E\{T_e\} \sim -\tfrac{1}{2}E\left(\frac{\partial^3 L}{\partial\theta_1\partial\theta_2^2}\right)\Big/(I_{22}\sqrt{I_{11}}). \tag{6}$$

For reference, the generalizations of (3) and (5) in the case of several nuisance parameters θ_i $(i = 2, \dots, m)$ are

$$\begin{aligned} T_e - T \sim {} & I^{ik}\frac{\partial L}{\partial\theta_i}\left[A\frac{\partial^2 L}{\partial\theta_1\partial\theta_k} + B_j\frac{\partial^2 L}{\partial\theta_j\partial\theta_k} + \frac{\partial L}{\partial\theta_1}\frac{\partial A}{\partial\theta_k} + \frac{\partial L}{\partial\theta_j}\frac{\partial B_j}{\partial\theta_k}\right] \\ & + \tfrac{1}{2}I^{ik}\frac{\partial L}{\partial\theta_i}\frac{\partial L}{\partial\theta_g}I^{gh}\left[AE\left(\frac{\partial^3 L}{\partial\theta_1\partial\theta_k\partial\theta_h}\right) + B_jE\left(\frac{\partial^3 L}{\partial\theta_j\partial\theta_k\partial\theta_h}\right) - 2I_{1k}\frac{\partial A}{\partial\theta_h} - 2I_{jk}\frac{\partial B_j}{\partial\theta_h}\right], \end{aligned} \tag{3a}$$

where
$$T \equiv A\frac{\partial L}{\partial \theta_1} + B_j \frac{\partial L}{\partial \theta_j},$$
say, and
$$E\{T_e\} \sim -\tfrac{1}{2}A\left[E\left\{\frac{\partial^3 L}{\partial\theta_1\partial\theta_k\partial\theta_h}\right\} I^{kh} + 2I^{ik}\frac{\partial I_{1k}}{\partial\theta_i}\right] - \tfrac{1}{2}B_j\left[E\left\{\frac{\partial^3 L}{\partial\theta_j\partial\theta_k\partial\theta_h}\right\} I^{kh} + 2I^{ik}\frac{\partial I_{jk}}{\partial\theta_i}\right], \quad (5a)$$
the summation convention being understood for $g, h, i, j, k\,(2, \ldots, m)$, and I^{ij} being the inverse of I_{ij}.

Provided T_e is corrected for its bias, it is found further from the formulae (3) and (3*a*) that its second and third moments are then the same as those of T to $O(1/\sqrt{n})$.

Examples. (i) and (ii). The two examples discussed in §§ 6 and 7 of II are not affected by this correction. In the analysis of variance problem of § 6, $I_{12} \neq 0$, but it will be found that $E(T_e) \sim 0$ from (5) above. (I have learnt that more direct investigations of the asymptotic confidence interval for this problem, using the methods first developed by Dr B. L. Welch in the Behrens–Fisher problem, have recently* been made by J. R. Green and by A. Huitson; and, in particular, Dr Huitson has checked the consistency of my solution with his own results.)

Similarly, no bias correction arises in the time-series problem of § 7, though there is unfortunately a further misstatement in equation (23), which should be corrected. The maximum-likelihood estimate of $\sigma_1^2 \equiv \alpha$, say, should have been given as
$$\hat{\alpha} = \sum_{r=1}^{n} (X_r - \beta X_{r-1})^2/n. \quad (7)$$
The confidence interval from (24) then becomes
$$\beta = \hat{\beta} \pm \frac{\mu}{\sqrt{n}}\sqrt{(1-\hat{\beta}^2)} + \frac{\hat{\beta}}{n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right) \quad (8)$$
(cf. my contribution to the discussion at the Royal Statistical Society Symposium on Interval Estimation, 12 May 1954).

(iii) To illustrate the use of the bias correction where it is not zero, consider the last example above when the mean value m of X is not zero and requires estimation. The log likelihood function (apart from end-corrections of relative order $1/n$) is now
$$L = -\frac{n}{2}\log\alpha - \frac{1}{2}\sum_{r=1}^{n} Y_r^2/\alpha, \quad (9)$$
where
$$Y_r = (X_r - m) - \beta(X_{r-1} - m). \quad (10)$$
The L derivatives with respect to α and β are as before, with $\xi_r \equiv X_r - m$ in place of X_r. Also
$$\frac{\partial L}{\partial m} = \frac{(1-\beta)}{\alpha}\sum_{r=1}^{n}(\xi_r - \beta\xi_{r-1}),$$
$$\hat{m} \sim \sum_{r=1}^{n} X_r/n,$$
$$I_{mm} = (1-\beta)^2 n/\alpha, \quad I_{\alpha m} = 0, \quad I_{\beta m} = 0,$$
$$\frac{\partial^3 L}{\partial\beta\,\partial m^2} = 2n(1-\beta)/\alpha = E\left\{\frac{\partial^3 L}{\partial\beta\partial m^2}\right\}.$$
Hence the quantity
$$T_e \equiv \frac{\sum_{r=1}^{n}(X_r' - \beta X_{r-1}')\,X_{r-1}'/\hat{\alpha}'}{\sqrt{\{n/(1-\beta^2)\}}}, \quad (11)$$

* See references.

where $X'_r \equiv X_r - \hat{m}$, and $\hat{\alpha}'$ is like $\hat{\alpha}$ in (7) but with X'_r in place of X_r, has bias

$$E\{T_e\} \sim -\frac{1}{(1-\beta)\sqrt{I_{\beta\beta}}},$$

and the equation for β becomes

$$T_e + \frac{\sqrt{(1-\beta^2)}}{(1-\beta)\sqrt{n}} - \frac{\beta(\mu^2-1)}{\sqrt{\{(1-\beta^2)n\}}} = \pm\mu, \tag{12}$$

or

$$\beta = \hat{\beta} \pm \frac{\mu}{\sqrt{n}}\sqrt{(1-\hat{\beta}^2)} + \frac{1+2\hat{\beta}}{n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right) \tag{13}$$

in place of (8), where now

$$\hat{\beta} \equiv \sum_{r=1}^{n} X'_r X'_{r-1} \Big/ \sum_{r=1}^{n} (X'_{r-1})^2.$$

(iv) Finally, in view of the above amendments, it seemed advisable to check these results to $O(1/\sqrt{n})$ by means of a fairly complicated example for which the result is already known. A useful case is that of the classical correlation coefficient ρ, with the two standard deviations α and β as nuisance parameters. We take, apart from an additive constant,

$$L = -n\log\{\alpha\beta\sqrt{(1-\rho^2)}\} - \frac{n}{2(1-\rho^2)}\left[\frac{s_1^2}{\alpha^2} - \frac{2\rho r s_1 s_2}{\alpha\beta} + \frac{s_2^2}{\beta^2}\right], \tag{14}$$

where r, s_1, s_2 are the usual sample estimates of ρ, α and β respectively (the means for convenience are already eliminated). Then (cf. M. G. Kendall's *Advanced Theory of Statistics*, vol. 2, ex. 17.18)

$$\left.\begin{aligned} I_{\alpha\alpha} &= \frac{n(2-\rho^2)}{\alpha^2(1-\rho^2)}, \quad I_{\beta\beta} = \frac{n(2-\rho^2)}{\beta^2(1-\rho^2)}, \quad I_{\rho\rho} = \frac{n(1+\rho^2)}{(1-\rho^2)^2}, \\ I_{\alpha\beta} &= \frac{-n\rho^2}{\alpha\beta(1-\rho^2)}, \quad I_{\alpha\rho} = \frac{-n\rho}{\alpha(1-\rho^2)}, \quad I_{\beta\rho} = \frac{-n\rho}{\beta(1-\rho^2)}. \end{aligned}\right\} \tag{15}$$

It is found further from (14) that

$$\left.\begin{aligned} E\left|\frac{\partial^3 L}{\partial\alpha^3}\right| &= \frac{2n(5-2\rho^2)}{\alpha^3(1-\rho^2)}, & E\left|\frac{\partial^3 L}{\partial\alpha^2\partial\beta}\right| &= \frac{-2n\rho^2}{\alpha^2\beta(1-\rho^2)}, \\ E\left|\frac{\partial^3 L}{\partial\alpha^2\partial\rho}\right| &= \frac{-2n\rho(1-\rho^2)}{\alpha^2(1-\rho^2)^2}, & E\left|\frac{\partial^3 L}{\partial\alpha\,\partial\beta\,\partial\rho}\right| &= \frac{n\rho(1+\rho^2)}{\alpha\beta(1-\rho^2)^2}, \\ E\left|\frac{\partial^3 L}{\partial\alpha\,\partial\rho^2}\right| &= \frac{2n(1+\rho^2)}{\alpha(1-\rho^2)^2}, & E\left|\frac{\partial^3 L}{\partial\rho^3}\right| &= \frac{-4n\rho(3+\rho^2)}{(1-\rho^2)^3}, \end{aligned}\right\} \tag{16}$$

the expressions for the remaining quantities with α and β interchanged being obvious from symmetry.

The statistic T uncorrelated with $\partial L/\partial\alpha$ and $\partial L/\partial\beta$ is

$$T \equiv \left(\frac{\partial L}{\partial\rho} + \frac{\alpha\rho}{2(1-\rho^2)}\frac{\partial L}{\partial\alpha} + \frac{\beta\rho}{2(1-\rho^2)}\frac{\partial L}{\partial\beta}\right) \Big/ \sqrt{I_{\rho\rho.\alpha\beta}}, \tag{17}$$

where $I_{\rho\rho.\alpha\beta} = n/(1-\rho^2)^2$. From the above formulae, and the general expression for $E\left|\frac{\partial L}{\partial\theta_i}\frac{\partial L}{\partial\theta_j}\frac{\partial L}{\partial\theta_k}\right|$ in terms of $E\left|\frac{\partial^3 L}{\partial\theta_i\partial\theta_j\partial\theta_k}\right|$ and derivatives of I_{ij} given in II, it is found (after some algebra) that the skewness of T in (17) is zero to $O(1/\sqrt{n})$.

We now replace T by T_e, the maximum-likelihood estimates of α and β (ρ known) being

$$\hat{\alpha} = s_1\sqrt{\left(\frac{1-\rho r}{1-\rho^2}\right)}, \quad \hat{\beta} = s_2\sqrt{\left(\frac{1-\rho r}{1-\rho^2}\right)}. \tag{18}$$

T in (17) then reduces to

$$T_e \equiv \frac{\sqrt{n}}{1-\rho^2}\left(\frac{\partial L}{\partial \rho}\right)_e = \frac{\sqrt{n}(r-\rho)}{1-\rho r}; \tag{19}$$

the functional form for r in (19) should be noted. From the general bias formula (5a), we find (noting that $I^{\alpha\alpha} = \frac{1}{4}(2-\rho^2)\,\alpha^2/n$, $I^{\alpha\beta} = \frac{1}{4}\rho^2\alpha\beta/n$)

$$E(T_e) = \tfrac{1}{2}\rho/\sqrt{n}. \tag{20}$$

Hence our confidence interval is given by

$$\frac{\sqrt{n}(r-\rho)}{1-\rho r} - \frac{\rho}{2\sqrt{n}} = \pm\frac{\mu}{\sqrt{n}}, \tag{21}$$

which agrees with the known results

$$E(z) \sim \zeta + \tfrac{1}{2}\rho/n, \quad \sigma^2(z) \sim 1/n, \quad \gamma_1(z) \sim 0,$$

where $$z = \tfrac{1}{2}\log_e\frac{1+r}{1-r}, \quad \zeta = \tfrac{1}{2}\log_e\frac{1+\rho}{1-\rho}.$$

While the above method is not of course required for this example, it is of some interest that the result (21) has been obtained merely by straightforward differentiation of the log likelihood function (14). It is, however, apparent that the algebra involved in getting the next term in the expansion becomes in general so intractable that a more direct attack on individual problems is then usually the more promising.

REFERENCES

BARTLETT, M. S. (1953). Approximate confidence intervals. II. More than one unknown parameter. *Biometrika*, **40**, 306.

GREEN, J. R. (1954). A confidence interval for variance components. *Ann. Math. Statist.* **25**, 671.

HUITSON, A. (1954). Ph.D. Thesis, University of Leeds.

THE THEORY OF CORRELATION BETWEEN TWO CONTINUOUS VARIABLES WHEN ONE IS DICHOTOMIZED

By ROBERT F. TATE

University of Washington†

1. Introduction and summary

The problem of biserial correlation arises when one is sampling from a bivariate normal population in which one of the variables has been dichotomized, giving rise to only two observable values, say 0 and 1, and one wishes to use this dichotomized sample to estimate, or to test hypotheses concerning, the correlation coefficient ρ of the original bivariate normal distribution. The problem of biserial correlation occurs frequently in psychological work, especially in test construction and validation.

The term biserial correlation was introduced by Karl Pearson (1909), who was the first to perceive the statistical importance of this particular type of problem. He proposed as an estimator the sample biserial correlation coefficient. The asymptotic variance of this estimator was derived by Soper (1913). Much literature exists on the subject of how best to compute Pearson's coefficient. In this connexion the reader should see Du Bois (1942), Dunlap (1936) and Royer (1941).

Prof. Harold Hotelling realized some years ago that the existing methods for dealing with the problem of biserial correlation were far from satisfactory, and suggested to the author that the whole situation be reconsidered. The results of this examination are contained in the present paper.

§ 2 contains a list of most of the notation which has been adopted, and § 3 deals with the mathematical model. In § 4 the question of maximum likelihood is treated. Asymptotic variances are derived for the estimators $\hat{\omega}$ and $\hat{\rho}$. The asymptotic variance for $\hat{\rho}$ is compared with the approximate expression arrived at by Maritz (1953) when he considered a somewhat restricted model. Both expressions are shown to achieve their minimum value at $\omega = 0$ when ρ is fixed.

Matters concerning asymptotic normality and asymptotic efficiency are also considered.

An appraisal of r^*, the sample biserial correlation coefficient, is given in detail in § 5. It is shown to have asymptotic efficiency for estimating ρ which is 1 when $\rho = 0$, but which approaches 0 when $|\rho|$ approaches 1. The well-known fact that r^* may be greater than 1 is pointed out and some notion of the magnitude of r^* is obtained by a consideration of the product-moment correlation coefficient r. Asymptotic normality of r^* is verified by the use of a theorem of Cramér. The asymptotic standard deviation is tabulated at the end of the paper (Table 2). A proof is given for the customarily assumed fact that the asymptotic variance has a minimum for fixed ρ when $\omega = 0$. For the case $\omega = 0$ an approximate variance stabilizing transformation is derived. Calculations pertaining to this transformation may

† Part of this research was done under an Office of Naval Research contract at the Institute of Statistics, University of North Carolina. The balance was sponsored by the Office of Naval Research on the Navy Theoretical Statistics Project at the Laboratory of Statistical Research, University of Washington.

be carried out by using a table (see Fisher, 1946, Table V B) for the function $\tanh^{-1} r$. This result should prove useful in many situations.

§ 6 is devoted to a discussion of an iterative method of solution for the likelihood equations. The method is essentially Newton's method for two variables, the calculated values ω^*, r^* being used to start the iteration. The computations are not really prohibitive, considering the importance of the problem, and are to a certain extent organizable for punched-cards methods. An example is given with all of the calculations illustrated. Values of $\phi(x)$, the reciprocal of Mills's ratio, are required for the solution of the likelihood equations. These may be obtained from the tables published as a separate contribution, immediately following the present paper.

Two matters of some importance which are not considered in the present paper are:

(1) An investigation of the bias of $\hat{\rho}$.

(2) The numerical tabulation of the asymptotic variances of $\hat{\omega}$ and $\hat{\rho}$.

Further study is indicated on at least the second point.

2. Notation

To eliminate the distraction of searching through the text, we shall list here most of the symbols and notational devices used:

$\psi(x,y) = \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\right]$, the bivariate normal density.

$\lambda(x) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2}$, the normal density.†

$p(x) = \int_x^{+\infty} \lambda(t)\,dt, \quad q(x) = 1-p(x).$†

$\phi(x) = \frac{\lambda(x)}{p(x)}$, the reciprocal of Mills's ratio.†

$\xi(x,\omega) = \int_\omega^{+\infty} \psi(x,y)\,dy.$

$\eta(x,\omega) = \int_{-\infty}^{\omega} \psi(x,y)\,dy.$

X the undichotomized normal random variable.

Y the dichotomized normal random variable.

ω the point of dichotomy of Y, measured in standard units.

Z the discrete random variable induced by the dichotomization of Y.

$f(x,z)$ the joint density of the random variables X and Z.

r^* the sample biserial correlation coefficient.

$AV(r^*)$ the asymptotic variance of r^*.

$AEff(r^*)$ the asymptotic efficiency of r^* for estimating ρ.

$\mathcal{N}(\mu,\sigma^2)$ a normal random variable with mean μ and variance σ^2.

† [*Editorial Note.* To bring this notation into conformity with that of the tables printed on pp. 217–221 below and with that used in the recently published *Biometrika Tables for Statisticians*, vol. 1, it is necessary to write $Z = \lambda(x)$, $Q = p(x)$, $P = 1-p(x)$, so that $\phi(x) = Z/Q$ or Z/P according as x is >0 or <0.]

3. MATHEMATICAL MODEL

Let (X, Y) have the bivariate normal distributon $\psi\{(x-\mu)/\sigma, (y-\nu)/\tau\}$. Let Z be a dichotomy of Y, with the point of dichotomy ω measured in standard units. Without losing any generality we may set $\nu = 0$ and $\tau = 1$. Z is thus a random variable which takes the value 1 when $Y \geqslant \omega$ and the value 0 when $Y < \omega$. Obviously,

$$P(Z=1) = \int_{\omega}^{+\infty} \lambda(y)\,dy = p(\omega), \quad P(Z=0) = q(\omega). \tag{3·1}$$

Consider a sample of n independent random vectors (X_1, Z_1), (X_2, Z_2), ..., (X_n, Z_n). The problem of biserial correlation consists of finding a suitable function of (X_i, Z_i) $(i = 1, 2, \ldots, n)$ with which to estimate ρ.

Karl Pearson (1909) introduced the estimator r^* ('biserial r'), which we express in the following form:†

$$r^* = \frac{\frac{1}{n}\Sigma(X_i-\bar{X})(Z_i-\bar{Z})}{\left\{\frac{1}{n}\Sigma(X_i-X)^2\right\}^{\frac{1}{2}}\lambda(T)} = r\frac{\left\{\frac{1}{n}\Sigma(Z_i-\bar{Z})^2\right\}^{\frac{1}{2}}}{\lambda(T)}, \tag{3·2}$$

where r is the product-moment correlation coefficient of (X_i, Z_i), and T is the solution of the equation‡

$$\int_{T}^{+\infty} \lambda(y)\,dy = \bar{Z}. \tag{3·3}$$

r^* will be discussed completely in § 5. For the present we shall merely state the asymptotic variance obtained by Soper (1913):

$$AV(r^*) = \frac{1}{n}\left\{\rho^4 + \rho^2\left[\frac{pq\omega^2}{\lambda^2} + \frac{(2p-1)\omega}{\lambda} - \frac{5}{2}\right] + \frac{pq}{\lambda^2}\right\}, \tag{3·4}$$

where the functions p, q and λ all have argument ω. $\sqrt{\{Av(r^*)\}}$ is given in Table 2 at the end of the paper. In view of symmetry about the values $\rho = 0$ and $p(\omega) = \frac{1}{2}$, the tabulation is given for $\rho = 0$ to 1 in steps of 0·10, and for $p = 0{\cdot}05$ to 0·50 in steps of 0·05.

Since the random variable Z takes the value 0 or 1, the joint density of (X, Z) can be written

$$f(x,z) = zf(x,1) + (1-z)f(x,0), \tag{3·5}$$

where

$$f(x,0) = \int_{-\infty}^{\omega} \Psi(x,y)\,dy, \quad f(x,1) = \int_{\omega}^{+\infty} \Psi(x,y)\,dy, \tag{3·6}$$

with $\Psi(x, y)$ denoting the bivariate normal density, $\psi\{(x-\mu)/\sigma, y\}$, with means μ and 0, variances σ^2 and 1, and correlation ρ. §§ 4 and 6 are devoted to a discussion of the likelihood function $\Pi f(x_i, z_i)$.

4. PROPERTIES OF THE MAXIMUM-LIKELIHOOD ESTIMATORS

As the likelihood function stands it may be expressed as

$$L(\mu, \sigma^2, \omega, \rho) = \Pi\left\{z_i\xi\left(\frac{x_i-\mu}{\sigma}, \omega\right) + (1-z_i)\eta\left(\frac{x_i-\mu}{\sigma}, \omega\right)\right\}. \tag{4·1}$$

Maritz (1953) considered the restricted model with $\mu = 0$, $\sigma^2 = 1$. Using biserial data (X_i, Z_i) $(i = 1, 2, \ldots, n)$, he introduces a grouping of the X observations, and then considers

† All Σ and Π symbols with index i will have limits 1 to n.

‡ The reason for this definition of T will be apparent in § 5, when we show that r^* is consistent and asymptotically normal.

the observations to be concentrated at their respective cell mid-points. This leads to a neat solution of the problem by probit-analysis methods. A proof of the convergence of this method as the grouping becomes finer must depend on a close examination of the limiting processes involved. Specifically, it is necessary that as the cell width becomes small, and the sample size large, each cell must contain sufficiently many observations that the ratio of the number of X observations whose corresponding Z observations are 1 to the number of X observations provides a valid approximation to the conditional probability that $Z = 1$. Instead of attempting a discussion of this point, we shall derive the asymptotic variances for the four-parameter problem (and, as a by-product, for the two-parameter problem), and in § 6 discuss an iterative method for obtaining $\hat{\omega}$ and $\hat{\rho}$. This method, while more time-consuming than that of Maritz, does not require grouping. It should be noted in this connexion that Tocher's exact method (see Tocher, 1949, pp. 9–11), also known as the 'scoring' method, does not help in this case, owing to the difficulty of obtaining expected second partial derivatives of L.

The likelihood situation of Maritz differs from ours because of the fact that when μ and σ^2 are set equal to 0 and 1 respectively, the customary method for obtaining asymptotic variances of $\hat{\omega}$ and $\hat{\rho}$ by an inversion of the information matrix leads to smaller variances. As far as the solution for $\hat{\omega}$ and $\hat{\rho}$ is concerned, the results will remain the same after a slight transformation.†

It may be remarked without dwelling at any length on regularity conditions that those given by Cramér (1946, p. 500) may be easily verified, since $f(x, 0)$ and $f(x, 1)$ are both integrals of bivariate normal densities. Consequently, $\hat{\mu}$, $\hat{\sigma}^2$, $\hat{\omega}$ and $\hat{\rho}$ will be asymptotically normal and asymptotically efficient estimators of the corresponding parameters. We now use the information matrix technique to find the asymptotic variances of $\hat{\omega}$ and $\hat{\rho}$.

THEOREM I. *The asymptotic variances of $\hat{\omega}$ and $\hat{\rho}$ are*

$$AV(\hat{\omega}) = \frac{1-\rho^2}{n}\left\{\frac{\int_{-\infty}^{+\infty}(x-\rho\omega)^2 g\,dx}{\int_{-\infty}^{+\infty} g\,dx\int_{-\infty}^{+\infty} x^2 g\,dx - \left(\int_{-\infty}^{+\infty} xg\,dx\right)^2}\right\} + \frac{\rho^2(\rho^2\omega^2+2)}{n},$$

$$AV(\hat{\rho}) = \frac{(1-\rho^2)^3}{n}\left\{\frac{\int_{-\infty}^{+\infty} g\,dx}{\int_{-\infty}^{+\infty} g\,dx\int_{-\infty}^{+\infty} x^2 g\,dx - \left(\int_{-\infty}^{+\infty} xg\,dx\right)^2}\right\} + \frac{\rho^2(1-\rho^2)^2}{n},$$

where

$$g(x,\omega,\rho) = \lambda(x)\,\phi\left(\frac{\omega-\rho x}{\sqrt{(1-\rho^2)}}\right)\phi\left(-\frac{\omega-\rho x}{\sqrt{(1-\rho^2)}}\right).$$

Proof. Using expression (4·1) for L, and letting δ^2 refer to any of the ten second order partial operators, we obtain the fundamental relation

$$E(\delta^2 \log L) = nqE_0(\delta^2 \log \eta) + npE_1(\delta^2 \log \xi), \tag{4·2}$$

where E_0 and E_1 mean conditional expectation with respect to the conditional densities of X given $Y < \omega$ and $Y \geqslant \omega$ respectively.

† The author expresses his indebtedness to the referee for this fact, and for its proof which will be given later.

For each of the possible operators δ^2 the calculation of (4·2) proceeds in about the same way. We compute, as an illustration,

$$E\left(\frac{\partial^2 \log L}{\partial \omega^2}\right).$$

It may be shown that

$$\Psi(x \mid Y < \omega) = \frac{1}{q}\eta\left(\frac{x-\mu}{\sigma}, \omega\right), \quad \Psi(x \mid Y \geqslant \omega) = \frac{1}{p}\xi\left(\frac{x-\mu}{\sigma}, \omega\right).$$

After some differentiation we get

$$\left.\begin{aligned} nqE_0\left(\frac{\partial^2 \log \eta}{\partial \omega^2}\right) &= -n\int_{-\infty}^{+\infty} \frac{\Psi^2(x,\omega)}{\eta\{(x-\mu)/\sigma, \omega\}}\,dx, \\ npE_1\left(\frac{\partial^2 \log \xi}{\partial \omega^2}\right) &= -n\int_{-\infty}^{+\infty} \frac{\Psi^2(x,\omega)}{\xi\{(x-\mu)/\sigma, \omega\}}\,dx. \end{aligned}\right\} \tag{4·3}$$

Combining the terms of (4·3) according to (4·2), making use of the relation

$$\int_{-\infty}^{+\infty} \Psi^2(x,\omega)\left\{\frac{1}{\eta\left(\frac{x-\mu}{\sigma}, \omega\right)} + \frac{1}{\xi\left(\frac{x-\mu}{\sigma}, \omega\right)}\right\} dx = \int_{-\infty}^{+\infty} g(x,\omega,\rho)\,dx,$$

and performing similar computations for the expectations of the other second partials of $\log L$, we arrive at the information matrix†

$$\begin{bmatrix} -\frac{a_0}{1-\rho^2} & \frac{a_1-\rho\omega a_0}{(1-\rho^2)^2} & -\frac{\rho a_0}{\sigma(1-\rho^2)} & -\frac{\rho a_1}{\sigma(1-\rho^2)} \\ & -\frac{a_2-2\rho\omega a_1+\rho^2\omega^2 a_0}{(1-\rho^2)^3} & \frac{a_1\rho-\rho^2\omega a_0}{\sigma(1-\rho^2)^2} & \frac{a_2\rho-\rho^2\omega a_1}{\sigma(1-\rho^2)^2} \\ \text{symmetric} & & -\frac{1-\rho^2+\rho^2 a_0}{\sigma^2(1-\rho^2)} & -\frac{\rho^2 a_1}{\sigma^2(1-\rho^2)} \\ & & & -\frac{2(1-\rho^2)+\rho^2 a_2}{\sigma^2(1-\rho^2)} \end{bmatrix},$$

where

$$a_k = \int_{-\infty}^{+\infty} x^k g(x,\omega,\rho)\,dx. \tag{4·4}$$

The asymptotic variances of $\hat{\omega}$ and $\hat{\rho}$, obtained after inversion, correspond with the expressions of the theorem.

The two-parameter problem of Maritz, solved by considering the upper left 2×2 submatrix, yields asymptotic variances which are the leading terms of the expressions given in the theorem; they coincide with his results.

The role played by ω in $AV(\hat{\rho})$ is partially described by the following theorem:

THEOREM II. *$AV(\hat{\rho})$ is a minimum for each ρ when $\omega = 0$.*

Proof. For the case $\rho = 0$ the proof follows from Theorem IV of the next section. For the case $\rho \neq 0$, make the transformation $(\omega - \rho x)(1-\rho^2)^{-\frac{1}{2}} = y$ in the integrand of (4·4). $AV(\hat{\rho})$ can then be expressed as

$$AV(\hat{\rho}) = \frac{\rho^2(1-\rho^2)^2}{n}\left\{1 + \frac{1}{b_2 - b_1^2/b_0}\right\},$$

where $b_k = E'[X^k \phi(X)\,\phi(-X)]$, with E' denoting expectation with respect to the distribution of $\mathcal{N}[\omega(1-\rho^2)^{-\frac{1}{2}}, \rho^2(1-\rho^2)^{-1}]$. It can now be seen from considerations of symmetry

† Columns (from left to right) and rows (from top to bottom) correspond to ω, ρ, μ, σ.

that b_0 and b_2 have maxima at $\omega = 0$, while b_1 has its minimum at $\omega = 0$, which completes the proof.

5. The utility of r^*

We now present a series of results concerning r^*, which will be followed by a general discussion of its value. Note that expression (3·2) for r^* is invariant under a linear transformation of the X_i, so in all results pertaining only to r^*, we set $\omega = 0$ and $\sigma^2 = 1$.

A little later we shall need $E(XZ)$. Since it is not difficult to obtain, we will give the expression for the general moment $\alpha_k = E(X^kZ)$.

Theorem III.
$$\alpha_k = \sum_{j=0}^{k} \binom{k}{j} (1-\rho^2)^{\frac{1}{2}j} \rho^{k-j} a_j \int_{\omega}^{+\infty} y^{k-j} e^{-\frac{1}{2}y^2} dy,$$

where a_j is the j-th moment of the random variable $\mathcal{N}(0, 1)$.

Proof. Using the definition of E_1, we obtain
$$E(X^kZ) = pE_1(X^k) = \int_{-\infty}^{+\infty}\int_{\omega}^{+\infty} x^k \psi(x, y)\, dy\, dx.$$
Make the transformation $t = (x - \rho y)\sqrt{(1-\rho^2)}$. The above then reduces to
$$\int_{-\infty}^{+\infty}\int_{\omega}^{+\infty} \{t\sqrt{(1-\rho^2)} + y\rho\}^k \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}y^2} dy\, dt.$$
Using a binomial expansion and integrating with respect to t, we obtain Theorem III. The integrals which occur are incomplete gamma functions which may be evaluated by the usual recursion relation.

For completeness we include the relation between $\rho(X, Y)$ and $\rho(X, Z)$, known to Karl Pearson:
$$\rho(X, Z) = \rho(X, Y) \frac{\lambda(\omega)}{\sqrt{(pq)}}.$$

It follows from the original definition of biserial correlation, as given by Pearson, that r^* is consistent. This fact is also an immediate consequence of relation (3·2) between r^* and r: $r \to \rho(X, Z)$ in probability as $n \to \infty$. Thus
$$r^* = \frac{r}{\lambda(T)} \left\{\frac{1}{n}\Sigma(Z_i - \bar{Z})^2\right\}^{\frac{1}{2}} \to \frac{\rho(X, Z)\sqrt{(pq)}}{\lambda(\omega)} \quad \text{in probability,}$$
and hence by the above, $r^* \to \rho(X, Y)$ in probability.

With respect to the magnitude of r^*, it is well known that $|r^*|$ can be greater than 1. Something of the nature of this phenomenon can be understood by looking at r. In order to prove a result concerning the magnitude of r^*, we shall need a preliminary result (see Tate, 1953, Lemma 2):

Theorem IV. $p(x)q(x) \geqslant \frac{1}{2}\pi\lambda^2(x)$, $(-\infty < x < +\infty)$, *with equality at* $0, \pm\infty$.

Now we have

Theorem V.
$$\frac{r^*}{r} \geqslant \sqrt{(\tfrac{1}{2}\pi)}.$$

Proof. Rewriting (3·2) as
$$r^* = r\frac{\sqrt{(\bar{Z} - \bar{Z}^2)}}{\lambda(T)},$$

we have, in view of the definition of T,

$$r^* = r\frac{\sqrt{\{p(T)\,q(T)\}}}{\lambda(T)}.$$

Theorem IV applies for any T, so Theorem V is proved. As a consequence of Theorem V, we see that

$$r^* \begin{matrix} > +1 \\ < -1 \end{matrix} \quad \text{according as} \quad r \begin{matrix} > +\sqrt{\frac{2}{\pi}} \\ < -\sqrt{\frac{2}{\pi}}. \end{matrix}$$

Asymptotic normality of r^*, which will be needed later in this section, is a consequence of a theorem of Cramér.

THEOREM VI. $$r^* \sim \mathcal{N}\{\rho, AV(r^*)\}.$$

Proof. In expression (3·2) the term $\lambda(T)$ is seen to be an infinitely differentiable function of $\bar{Z}$. Thus, r^* is a totally differentiable function of the sample means $\bar{X}, \bar{Z}, \bar{X^2}, \overline{XZ}$. Applying Cramér's theorem (see Cramér, 1946, p. 366), we have asymptotic normality with the asymptotic variance (3·4) calculated by Soper.

We shall now present two results which are more important than those just preceding. They concern the asymptotic, or large-sample, efficiency of r^*, with respect to the class of estimators of ρ based on the sample (X_i, Z_i).

THEOREM VII. *r^* is an asymptotically* most *efficient estimator of ρ when $\rho = 0$.*

Proof. In view of Theorem VI on asymptotic normality, we have a right to inquire about the asymptotic efficiency of r^*, which will be denoted by

$$AEff(r^*) = \frac{AV(\hat{\rho})}{AV(r^*)}.$$

It may be seen from Theorem I that

$$AV(\hat{\rho} \mid \omega, 0) = \frac{p(\omega)\,q(\omega)}{n\{\lambda(\omega)\}^2}. \tag{5·1}$$

Now, from (3·4) we observe that (5·1) coincides with $AV(r^* \mid \omega, 0)$. The conclusion follows from the definition of an asymptotically most efficient estimator.

THEOREM VIII. *r^* is an asymptotically* least *efficient estimator of ρ when $|\rho| \to 1$.*

Proof. An application of Theorem IV shows that

$$\phi\left(\frac{\omega - \rho x}{\sqrt{(1-\rho^2)}}\right)\phi\left(-\frac{\omega - \rho x}{\sqrt{(1-\rho^2)}}\right) \leqslant \frac{2}{\pi}.$$

Hence, recalling the definition of $g(x, \omega, \rho)$ in Theorem I, we see that all integrals of the form $\int_{-\infty}^{+\infty} x^k g(x, \omega, \rho)\,dx$ exist. Schwarz's inequality shows that $AV(\hat{\rho} \mid \omega, \rho)$ is such that the term in braces is non-vanishing. Thus, $AV(\hat{\rho} \mid \omega, \rho) \to 0$ as $|\rho| \to 1$. From the fact $AV(r^* \mid \omega, \rho) \to \frac{2}{\pi}$ as $|\rho| \to 1$, we conclude that $AEff(r^* \mid \omega, \rho) \to 0$.

The special case $\omega = 0$ has interesting features which will appear in Theorems X and XI. First we shall need another preliminary result (see Tate, 1953, Lemma 1):

THEOREM IX. $$\{1 - 2p(x)\}\,\lambda(x) - xp(x)\,q(x) \geqslant 0, \quad (x \geqslant 0).$$

THEOREM X. *The asymptotic variance of r^* has its minimum for each ρ at $\omega = 0$.*

Proof. In view of symmetry, it will be sufficient to show the result for $\omega \geqslant 0$. Let

$$A(\omega) = \frac{\omega^2 p(\omega)\, q(\omega)}{\{\lambda(\omega)\}^2} - \frac{\omega\{1-2p(\omega)\}}{\lambda(\omega)}, \quad B(\omega) = \frac{p(\omega)\, q(\omega)}{\{\lambda(\omega)\}^2},$$

$$g(\omega) = \{1-2p(\omega)\}\lambda(\omega) - \omega p(\omega)\, q(\omega), \quad h(\omega) = p(\omega)\, q(\omega) - \pi\{\lambda(\omega)\}^2/2.$$

From this point until the end of the proof, we shall omit ω whenever it appears as an argument of any function. We have (Tate, 1953)

$$g' = 2\lambda^2 - \pi\omega\lambda, \quad g = -4\omega\lambda^2 + (2q-1)\lambda, \quad g(0) = g(\infty) = 0, \quad g'(0) > 0,$$

$$h' = \lambda(1 - \omega\lambda), \quad h = \pi\omega\lambda^2, \quad h'' = \lambda^2(\pi - 2 - 2\pi\omega^2) - \omega(1-2q)\lambda,$$

$$h(0) = h(\infty) = h'(0) = 0, \quad h''(0) > 0.$$

Accordingly, we have

$$A = -\omega g \lambda^{-2}, \quad B = h\lambda^{-2} + \tfrac{1}{2}\pi, \quad \text{with} \quad A \leqslant 0, \quad B \geqslant \tfrac{1}{2}\pi,$$

both equalities holding at $\omega = 0$. The relation $AV(r^* \mid \omega, \rho) \geqslant AV(r^* \mid 0, \rho)$ for all ρ may be written $\rho^2 A + B \geqslant \frac{1}{2}\pi$ for all ρ. Since $A \leqslant 0$ this last expression is implied by $A + B \geqslant \frac{1}{2}\pi$, which in turn is equivalent to $h \geqslant \omega g$. Thus, we must show $k = h - \omega g \geqslant 0$:

$$\left.\begin{aligned} k' &= 2\omega q(1-q) - 2(2q-1)\lambda + \omega(\pi-2)\lambda^2, \\ k'' &= 2q(1-q) - \lambda^2\{6 - \pi + \omega^2(2\pi-4)\}, \\ k(0) = k(\infty) &= k'(0) = 0, \quad k''(0) = 1 - 3/\pi > 0. \end{aligned}\right\} \tag{5·2}$$

We shall show that there exists no y such that $k'(y) = 0$, $k(y) < 0$. Suppose such a y does exist. Then

$$\left.\begin{aligned} 2(2q-1)\lambda &= 2yq(1-q) + (\pi-2)y\lambda^2, \\ q(1-q)(1+y^2) - \tfrac{1}{2}\pi\lambda^2 - y(2q-1)\lambda &< 0. \end{aligned}\right\} \tag{5·3}$$

Substituting the right member of the first expression into the second, we have

$$2q(1-q) < \lambda^2\{\pi + (\pi-2)y^2\}.$$

Thus, $k''(y) < \lambda^2\{2\pi - 6 - (\pi-2)y^2\}$. A negative maximum must, from (5·2), be followed by a negative minimum. Hence, from the above relation in $k''(y)$, there exist no extrema which exceed $\{(2\pi-6)/(\pi-2)\}^{\frac{1}{2}}$. Assuming there is a negative extremum of k, then there must be a negative minimum in $(0, 1)$. Let y be this minimum point. Then $k''(y) > 0$, or from (5·2), $2q(1-q) - \lambda^2\{6 - \pi - (2\pi-4)y^2\} > 0$. Substituting the value of $2q(1-q)$ obtained from the first equation in (5·3), we reach $(2q-1) - y\lambda\{2 + (\pi-2)y^2\} > 0$. The left member vanishes at $y = 0$ and has a negative derivative for $0 \leqslant y^2 \leqslant 1$. Therefore, there is no negative minimum in $(0, 1)$, and from the previous argument $k \geqslant 0$, which completes the proof.

Since for any fixed ρ, r^* is a better estimator when $\omega = 0$, it will be useful to have for this case something simpler in the way of an asymptotic distribution of r^* than that contained in Theorem VI. We are therefore led to

THEOREM XI. *When $\omega = 0$, we have to a close approximation*

$$\tanh^{-1}\frac{2r^*}{\sqrt{5}} \sim \mathcal{N}\left(\tanh^{-1}\frac{2\rho}{\sqrt{5}}, \frac{5}{4n}\right).$$

Proof. $$AV(r^* \mid 0, \rho) = \frac{1}{n}(\rho^4 - \tfrac{5}{2}\rho^2 + \tfrac{1}{2}\pi) = \frac{1}{n}(\tfrac{5}{4} - \rho^2)^2 - \left(\frac{25 - 8\pi}{16n}\right).$$

Dropping the last term and solving the equation

$$\sqrt{g'(x)} = \frac{1}{(\frac{5}{4} - x^2)},$$

we have $g(x) = (2/\sqrt{5})\tanh^{-1}(2x/\sqrt{5})$. It is known that

$$\sqrt{n}\{g(r^*) - g(\rho)\} \sim \mathcal{N}(0, 1),$$

so the theorem is proved.

Discussion of results concerning r^*

In looking over Theorems V, VI, VII, VIII, X and XI, several facts stand out. First, even though r^* is consistent and asymptotically normal, it is still inadequate for estimating ρ because of its possible magnitude and its lack of large-sample efficiency for large values of $|\rho|$. In the case of testing the hypothesis $H: \rho = \rho_0$ the first defect is not of so much consequence. Even in a problem of estimation, one can always operate under the rule: when $|r^*| < 1$, estimate ρ by r^*; when $r^* > 1$, estimate $\rho = 1$; and when $r^* < -1$, estimate $\rho = -1$. The gross defect is lack of efficiency. In practically all applications it is of more interest to detect large values of ρ than small values. In just such cases r^* is a 'worst' estimator. On the other hand, again speaking in large-sample terms, when $\rho = 0$, r^* is a 'best' estimator.† Hence, if we base a test of $H: \rho = \rho_0$ on r^*, good results should be achieved when $|\rho_0|$ is small. It is then recommended that r^* be used for one and only one purpose, to test $H: \rho = \rho_0$ when $|\rho_0|$ is small. If, in addition, the assumption $\omega = 0$ is tenable, then the variance stabilizing transformation of Theorem XI may be used, calculations being performed with Table VB of Fisher (1946, p. 210). In such a case certain advantages (see Fisher 1946, pp. 197–204) will accrue. $\{nAV(r^*)\}^{\frac{1}{2}}$ is given in Table 2. Note that from Theorem X, $\omega = 0$ is desirable on the grounds of precision.

6. Solution of the likelihood equations

Using the notation of (4·1), we have the likelihood equations

$$\delta L = \Sigma \left\{ z_i \frac{\delta\xi\left(\frac{x_i - \mu}{\sigma}, \omega\right)}{\xi\left(\frac{x_i - \mu}{\sigma}, \omega\right)} + (1 - z_i) \frac{\delta\eta\left(\frac{x_i - \mu}{\sigma}, \omega\right)}{\eta\left(\frac{x_i - \mu}{\sigma}, \omega\right)} \right\} = 0. \tag{6·1}$$

As was mentioned previously (cf. footnote, p. 208), the solution of the four-parameter problem reduces to the problem of determining $\hat{\omega}$ and $\hat{\rho}$. It turns out that the likelihood equations for μ and ω may be combined to yield $\hat{\mu} = \bar{x}$. Similarly, a combination of likelihood equations for σ^2 and ρ will give us $\hat{\sigma}^2 = s^2$. Details will be omitted.

We now replace μ and σ by $\bar{x}$ and s in the expression $(x_i - \mu)/\sigma$, which occurs in (6·1), and denote the result by x_i'. Also, let L' denote the new likelihood function. The solution of $\delta L' = 0$ for $\hat{\omega}$ and $\hat{\rho}$ follows.

† Another point in favour of r^* is indicated by the parallelism of Theorems II and X.

One may easily verify that

$$\left.\begin{aligned} \frac{\partial\eta(x_i',\omega)}{\partial\omega} &= \psi(x_i',\omega), & \frac{\partial\xi(x_i',\omega)}{\partial\omega} &= -\psi(x_i',\omega), \\ \frac{\partial\eta(x_i',\omega)}{\partial\rho} &= \frac{-(x_i'-\rho\omega)\,\psi(x_i',\omega)}{(1-\rho^2)}, & \frac{\partial\xi(x_i',\omega)}{\partial\rho} &= \frac{(x_i'-\rho\omega)\,\psi(x_i',\omega)}{(1-\rho^2)}. \end{aligned}\right\} \tag{6·2}$$

By the use of (6·2) the equations $\delta L' = 0$ can be written as

$$\left.\begin{aligned} \Sigma(x_i'-\rho\omega)\,(2z_i-1)\,\phi\left\{(2z_i-1)\left(\frac{\omega-\rho x_i'}{\sqrt{(1-\rho^2)}}\right)\right\} &= 0, \\ \Sigma(2z_i-1)\,\phi\left\{(2z_i-1)\left(\frac{\omega-\rho x_i'}{\sqrt{(1-\rho^2)}}\right)\right\} &= 0. \end{aligned}\right\} \tag{6·3}$$

Now introduce the notation

$$\left.\begin{aligned} \delta_i = 2z_i-1, \quad \gamma_i = (\omega-\rho x_i')\,(1-\rho^2)^{-\frac{1}{2}}, \quad \phi_i = \phi(\delta_i\gamma_i), \\ A_i = \phi_i(\phi_i-\delta_i\gamma_i). \end{aligned}\right\} \tag{6·4}$$

Rewriting (6·3) again, in the new notation, we have

$$\Sigma\delta_i\phi_i = 0, \quad \Sigma\delta_i x_i'\phi_i = 0. \tag{6·5}$$

Easy differentiation gives $\phi'(x) = \phi(x)\{\phi(x)-x\}$. Newton's method in two variables gives the following equations in $\Delta\omega$ and $\Delta\rho$, where $\Delta\omega = \omega-\omega_1$, $\Delta\rho = \rho-\rho_1$, ω_1 and ρ_1 being initial guesses:

$$\left.\begin{aligned} \left(\Sigma\frac{A_i}{\sqrt{(1-\rho^2)}}\right)\Delta\omega + \left(\frac{\rho_1\omega_1\Sigma A_i - \Sigma A_i x_i'}{(1-\rho_1^2)^{\frac{1}{2}}}\right)\Delta\rho &= -\Sigma\delta_i\phi_i, \\ \left(\Sigma\frac{A_i x_i'}{\sqrt{(1-\rho^2)}}\right)\Delta\omega + \left(\frac{\rho_1\omega_1\Sigma A_i x_i' - \Sigma A_i x_i'^2}{(1-\rho_1^2)^{\frac{1}{2}}}\right)\Delta\rho &= -\Sigma\delta_i x_i'\phi_i. \end{aligned}\right\} \tag{6·6}$$

Let Δ be the determinant of the coefficients. The method of solution will then be the following:

(i) Compute ω^*, r^* from the sample (x_i, z_i) $(i = 1, 2, \ldots, n)$, where r^* is the sample biserial correlation coefficient and ω^* is the solution of the equation $p(\omega) = \bar{z}$. Now, let $\omega_1 = \omega^*$ and

$$\rho_1 = \begin{cases} r^* & \text{when} \quad |r^*| < 1, \\ +0{\cdot}90 & \text{when} \quad r^* \geqslant 1, \\ -0{\cdot}90 & \text{when} \quad r^* \leqslant -1. \end{cases}$$

(ii) Compute δ_i, γ_i, ϕ_i, $\delta_i\phi_i$, $\delta_i x_i'\phi_i$, A_i, $A_i x_i'$, $A_i x_i'^2$ for $i = 1, 2, \ldots, n$, where δ_i, γ_i, ϕ_i, A_i are defined in (6·4), and the numerical values of the $\phi_i = \phi(\delta_i\gamma_i)$ are obtained from the tables printed on pp. 217–221 below. Note that these tables must be entered with $X = \delta_i\gamma_i$, while $\phi_i = Z/Q$ if $X > 0$ and Z/P if $X < 0$.

(iii) Evaluate the three determinants

$$\Delta = \begin{vmatrix} \Sigma A_i & \rho_1\omega_1\Sigma A_i - \Sigma A_i x_i' \\ \Sigma A_i x_i' & \rho_1\omega_1\Sigma A_i x_i' - \Sigma A_i x_i'^2 \end{vmatrix} \frac{1}{(1-\rho_1^2)^2},$$

$$\Delta\omega = \begin{vmatrix} -\Sigma\delta_i\phi_i & \rho_1\omega_1\Sigma A_i - \Sigma A_i x_i' \\ -\Sigma\delta_i x_i'\phi_i & \rho_1\omega_1\Sigma A_i x_i' - \Sigma A_i x_i'^2 \end{vmatrix} \frac{1}{\Delta(1-\rho_1^2)^{\frac{1}{2}}},$$

$$\Delta\rho = \begin{vmatrix} \Sigma A_i & -\Sigma\delta_i\phi_i \\ \Sigma A_i x_i' & -\Sigma\delta_i x_i'\phi_i \end{vmatrix} \frac{1}{\Delta(1-\rho_1^2)^{\frac{1}{2}}}.$$

(iv) Obtain (ω, ρ) from $(\Delta\omega, \Delta\rho)$ and (ω_1, ρ_1), and repeat the process using $\omega = \omega_2$, $\rho = \rho_2$ in place of ω_1 and ρ_1.

The rule given in (i) is somewhat arbitrary, but is believed to be a good rule of thumb. The longest stage in the scheme outlined above is the determination of ϕ_i, $i = 1, 2, \ldots, n$, from the tables.

We shall now present an illustration of the method. In order to have a good vantage point for observing the way the calculations run, we select a random sample from a bivariate normal population with zero means, unit variances and correlation $1/\sqrt{2}$. A table of random numbers from such a population is not available directly†, but can be constructed from a table of random numbers from $\mathcal{N}(0, 1)$ as follows: Let $U = \mathcal{N}(0, 1)$, $V = \mathcal{N}(0, 1)$, $\omega = \frac{1}{2}$, with U and V independent. Now let $X = U$ and $Y = (U + V)/\sqrt{2}$. After introducing the dichotomy

$$Z = 1 \text{ if } Y \geqslant \tfrac{1}{2}, \quad Z = 0 \text{ if } Y < \tfrac{1}{2}, \tag{6·7}$$

we have the desired set-up. It is important to note that for this particular example, the x_i' of the computing scheme become merely x_i, since it is known that $\mu = 0$ and $\sigma^2 = 1$. For computations see Table 1 below.

A second iteration resulted in $\omega_3 = 0{\cdot}251$, $\rho_3 = 0{\cdot}489$. Since $\hat{\rho}$ remained unchanged in the third decimal, the results were not included. Recall that the true value of $\hat{\rho}$ is 0·707. On the basis of our sample of 20, $\hat{\rho} = 0{\cdot}489$ is the best we can do. However, by using the iterative scheme instead of r^* we removed 27 % of the error.

Although r^* is used merely to start the iteration, the fact that

$$|r^* - E(r^*)|/\sqrt{\{AV(r^*)\}} = 1{\cdot}64$$

serves to indicate that we could have been more fortunate in our selection of a sample.

REFERENCES

CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. Princeton: Princeton University Press.

DU BOIS, P. H. (1942). A note on the computation of biserial r in item validation. *Psychometrika*, **7**, 143.

DUNLAP, J. W. (1936). A nomograph for computing biserial r. *Psychometrika*, **1**, 59.

FISHER, R. A. (1946). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.

MARITZ, J. S. (1953). Estimation of the correlation coefficient of a bivariate normal population where one of the variables is dichotomized. *Psychometrika*, **18**, 97.

PEARSON, K. (1909). On a new method for determining the correlation between a measured character A and a character B. *Biometrika*, **7**, 96.

ROYER, E. B. (1941). Punched cards methods for determining biserial coefficients. *Psychometrika*, **6**, 55.

SOPER, H. E. (1913). On the probable error for the biserial expression for the correlation coefficient. *Biometrika*, **10**, 384.

TATE, R. F. (1953). On a double inequality of the normal distribution. *Ann. Math. Statist.* **24**, 133.

TOCHER, K. D. (1949). A note on the analysis of grouped probit data. *Biometrika*, **36**, 9.

† [A table of *Correlated Random Normal Deviates* is now at press and will be issued shortly from the Department of Statistics, University College, London as *Tracts for Computers* No. XXVI, ED.]

Table 1. *Computation of example*

i	x_i	z_i	γ_i	δ_i	ϕ_i	$\delta_i x_i \phi_i$	$\phi_i - \delta_i \gamma_i$	A_i	$A_i x_i$	$A_1 x_1^2$
1	0·24	0	0·030	—	0·779	0·187	0·809	0·630	0·151	0·036
2	0·63	1	−0·146		0·707	0·445	0·853	0·603	0·380	0·239
3	−0·59	0	0·404	—	0·560	−0·330	0·964	0·540	−0·319	0·188
4	1·08	0	−0·348	—	1·032	1·115	0·684	0·706	0·762	0·822
5	0·06	0	0·111	—	0·729	0·044	0·840	0·451	0·027	0·002
6	−0·01	0	0·143	—	0·709	−0·007	0·852	0·604	−0·006	0·000
7	1·59	1	−0·578		0·470	0·747	1·048	0·493	0·784	1·247
8	−0·41	0	0·323	—	0·604	−0·248	0·927	0·560	−0·230	0·094
9	−0·33	1	0·287		0·989	−0·326	0·702	0·694	−0·229	0·076
10	0·99	1	−0·308		0·613	0·607	0·921	0·565	0·559	0·553
11	0·30	0	0·003	—	0·796	0·239	0·799	0·636	0·191	0·057
12	−2·07	0	1·070	—	0·262	−0·542	1·332	0·349	−0·722	1·495
13	−0·21	0	0·233	—	0·655	−0·138	0·888	0·582	−0·122	0·026
14	−0·47	1	0·350		1·033	0·486	0·683	0·706	−0·332	0·156
15	1·28	1	−0·438		0·541	0·692	0·979	0·530	0·678	0·868
16	0·82	1	−0·231		0·656	0·538	0·887	0·582	0·477	0·391
17	1·06	1	−0·339		0·595	0·631	0·934	0·556	0·589	0·624
18	1·09	0	−0·353	—	1·035	1·128	0·682	0·706	0·770	0·829
19	0·57	0	−0·119	—	0·875	0·499	0·756	0·662	0·377	0·215
20	−0·53	0	0·377	—	1·052	−0·558	0·675	0·710	−0·376	0·199

$\Sigma x_i = 5{\cdot}09$ $\quad \Sigma x_i z_i = 5{\cdot}04$ $\quad r^* = +0{\cdot}410\,(\rho_1)$ $\quad \gamma_i = 0{\cdot}138 - 0{\cdot}450 x_i$

$\Sigma x_1^2 = 15{\cdot}626$ $\quad \Sigma z_i = 9$ $\quad \omega^* = +0{\cdot}126\,(\omega_1)$

$\Sigma \delta_i \phi_i = -1{\cdot}380$ $\quad \Sigma A_i = 11{\cdot}865$ $\quad \Sigma A_i x_1^2 = 8{\cdot}127$ $\quad \Delta = -122{\cdot}54$

$\Sigma \delta_i x_i \phi_i = 0{\cdot}343$ $\quad \Sigma A_i x_i = 3{\cdot}409$ $\quad \Delta\omega = 0{\cdot}108$ $\quad \Delta\rho = 0{\cdot}079$

$\omega_2 = 0{\cdot}234$ $\quad \rho_2 = 0{\cdot}489$

Table 2. *The asymptotic standard deviation of r^* (biserial r) as a function of p and ρ (equation (3·4))*

All values must be divided by $\sqrt{n}$, as the quantity tabled is $\{nA\,V(r^*)\}^{\frac{1}{2}}$

$\pm\rho$ \ p or $1-p$	0·05	0·10	0·15	0·20	0·25	0·30	0·35	0·40	0·45	0·50
0·00	4·466	2·922	2·345	2·041	1·857	1·737	1·658	1·608	1·580	1·571
0·10	2·104	1·699	1·521	1·419	1·353	1·308	1·278	1·258	1·247	1·243
0·20	2·077	1·668	1·491	1·389	1·323	1·279	1·248	1·228	1·217	1·213
0·30	2·033	1·616	1·440	1·339	1·273	1·229	1·198	1·179	1·167	1·163
0·40	1·971	1·543	1·370	1·269	1·203	1·159	1·128	1·109	1·097	1·093
0·50	1·893	1·449	1·279	1·179	1·114	1·069	1·038	1·019	1·008	1·004
0·60	1·799	1·333	1·167	1·069	1·004	0·960	0·930	0·910	0·898	0·894
0·70	1·691	1·194	1·034	0·939	0·875	0·831	0·801	0·781	0·769	0·766
0·80	1·569	1·031	0·881	0·789	0·727	0·683	0·653	0·632	0·620	0·616
0·90	1·438	0·842	0·705	0·619	0·559	0·517	0·486	0·465	0·453	0·449
1·00	1·302	0·616	0·503	0·429	0·374	0·335	0·304	0·283	0·270	0·266

THE NORMAL PROBABILITY FUNCTION: TABLES OF CERTAIN AREA-ORDINATE RATIOS AND OF THEIR RECIPROCALS

EDITORIAL

Notation. Writing X as a standardized normal deviate, the following notation will be used:*

$$Z = Z(X) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}X^2},$$

$$P = P(X) = \int_{-\infty}^{X} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}u^2}\, du, \quad Q = 1 - P = \int_{X}^{\infty} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}u^2}\, du.$$

The table then gives for $X = 0{\cdot}00(0{\cdot}01)3{\cdot}00$ the values of the four ratios P/Z, Q/Z, Z/P and Z/Q. For the last three ratios, the results are extended for argument $X = 3{\cdot}00(0{\cdot}01)4{\cdot}00(0{\cdot}05)5{\cdot}00$. The ratios P/Z, Q/Z and Z/Q are tabled to five decimal places and those for Z/P to five significant figures.

Derivation. The ratio P/Z was freshly computed in the Department of Statistics, University College, London, using the *Tables of Normal Probability Functions*, National Bureau of Standards (1953). The ratio Q/Z has been taken directly from Table III of Karl Pearson's *Tables for Statisticians and Biometricians*, Part II (1931). The ratios Z/P and Z/Q were calculated by Prof. Z. W. Birnbaum using the National Bureau of Standards Tables, and were originally included in two tables forming part of the preceding paper by Dr R. F. Tate (1955). With Dr Tate's and Prof. Birnbaum's approval it was, however, decided to give separately the more comprehensive table which follows. This will form a useful companion to Table II, *Tables for Statisticians and Biometricians*, Part II, which gives the same four ratios for an argument of P instead of X. It is intended to reproduce both tables in the second volume of *Biometrika Tables for Statisticians*.

Having regard to this wider use, Dr Tate's notation has been modified, so that his $\phi(x)$ (see p. 206 above) becomes Z/P for $x = -X < 0$ and Z/Q for $x = X > 0$.

Interpolation. For a large part of the table accuracy to the full number of decimal places given may be obtained by linear interpolation. Where this fails, the Bessel formula

$$y_\theta = (1-\theta)\, y_0 + \theta y_1 - \tfrac{1}{4}\theta(1-\theta)\, (\delta^2 y_0 + \delta^2 y_1)$$

is adequate, except for P/Z with $X > 2{\cdot}50$, when full use of second differences in Everett's formula becomes necessary.

The normal probability function; extension of main table

X	Q/Z	Z/P	Z/Q	X	Q/Z	Z/P	Z/Q
4·00	0·23665	$0{\cdot}0^3$13383	4·22561	**4·50**	0·21257	$0{\cdot}0^4$15984	4·70432
·05	·23401	$\cdot0^3$10944	4·27330	·55	·21042	$\cdot0^4$12747	4·75240
·10	·23143	$\cdot0^4$89264	4·32103	·60	·20831	$\cdot0^4$10141	4·80051
·15	·22890	$\cdot0^4$72627	4·36880	·65	·20624	$\cdot0^5$80472	4·84865
·20	·22642	$\cdot0^4$58944	4·41662	·70	·20421	$\cdot0^5$63698	4·89682
4·25	0·22399	$0{\cdot}0^4$47719	4·46447	**4·75**	0·20222	$0{\cdot}0^5$50295	4·94503
·30	·22161	$\cdot0^4$38536	4·51237	·80	·20027	$\cdot0^5$39613	4·99326
·35	·21928	$\cdot0^4$31042	4·56030	·85	·19835	$\cdot0^5$31122	5·04153
·40	·21700	$\cdot0^4$24943	4·60827	·90	·19647	$\cdot0^5$24390	5·08983
·45	·21476	$\cdot0^4$19992	4·65628	·95	·19462	$\cdot0^5$19066	5·13815
4·50	0·21257	$0{\cdot}0^4$15984	4·70432	**5·00**	0·19281	$0{\cdot}0^5$14867	5·18650

* This conforms to the notation used by Pearson and Hartley in *Biometrika Tables for Statisticians*, 1 (1954).

The normal probability function. Table of area/ordinate ratios and their reciprocals

X	P/Z	Q/Z	Z/P	Z/Q	X	P/Z	Q/Z	Z/P	Z/Q
0·00	1·25331	1·25331	0·79788	0·79788	0·50	1·96402	0·87636	0·50916	1·14108
·01	1·26338	1·24338	·79153	·80426	·51	1·98399	·87078	·50404	1·14840
·02	1·27357	1·23356	·78520	·81066	·52	2·00426	·86525	·49894	1·15574
·03	1·28389	1·22387	·77888	·81708	·53	2·02483	·85977	·49387	1·16310
·04	1·29434	1·21430	·77260	·82352	·54	2·04572	·85436	·48882	1·17047
0·05	1·30492	1·20484	0·76633	0·82999	0·55	2·06693	0·84900	0·48381	1·17786
·06	1·31564	1·19550	·76008	·83647	·56	2·08846	·84370	·47882	1·18526
·07	1·32650	1·18627	·75386	·84298	·57	2·11032	·83845	·47386	1·19268
·08	1·33750	1·17716	·74766	·84950	·58	2·13252	·83326	·46893	1·20011
·09	1·34864	1·16816	·74149	·85605	·59	2·15506	·82812	·46402	1·20756
0·10	1·35993	1·15926	0·73533	0·86262	0·60	2·17795	0·82303	0·45915	1·21503
·11	1·37136	1·15048	·72920	·86921	·61	2·20120	·81799	·45430	1·22251
·12	1·38295	1·14179	·72309	·87582	·62	2·22481	·81301	·44948	1·23000
·13	1·39468	1·13321	·71701	·88244	·63	2·24879	·80807	·44468	1·23751
·14	1·40658	1·12474	·71095	·88910	·64	2·27315	·80319	·43992	1·24504
0·15	1·41862	1·11636	0·70491	0·89577	0·65	2·29789	0·79835	0·43518	1·25258
·16	1·43083	1·10809	·69889	·90246	·66	2·32302	·79357	·43047	1·26013
·17	1·44320	1·09991	·69290	·90917	·67	2·34855	·78883	·42579	1·26770
·18	1·45574	1·09183	·68694	·91590	·68	2·37449	·78414	·42114	1·27529
·19	1·46844	1·08384	·68099	·92265	·69	2·40085	·77949	·41652	1·28289
0·20	1·48132	1·07594	0·67507	0·92942	0·70	2·42763	0·77489	0·41192	1·29050
·21	1·49437	1·06814	·66918	·93620	·71	2·45484	·77034	·40736	1·29813
·22	1·50760	1·06043	·66331	·94301	·72	2·48249	·76583	·40282	1·30577
·23	1·52101	1·05281	·65746	·94984	·73	2·51059	·76137	·39831	1·31342
·24	1·53460	1·04527	·65164	·95669	·74	2·53915	·75695	·39383	1·32109
0·25	1·54837	1·03782	0·64584	0·96355	0·75	2·56817	0·75257	0·38938	1·32878
·26	1·56234	1·03046	·64007	·97044	·76	2·59767	·74824	·38496	1·33648
·27	1·57650	1·02318	·63432	·97734	·77	2·62766	·74394	·38057	1·34419
·28	1·59085	1·01599	·62859	·98426	·78	2·65814	·73969	·37620	1·35191
·29	1·60541	1·00887	·62290	·99121	·79	2·68913	·73548	·37187	1·35965
0·30	1·62017	1·00184	0·61722	0·99817	0·80	2·72063	0·73131	0·36756	1·36740
·31	1·63513	0·99488	·61157	1·00514	·81	2·75266	·72718	·36328	1·37517
·32	1·65030	·98801	·60595	1·01214	·82	2·78523	·72309	·35904	1·38295
·33	1·66569	·98121	·60035	1·01916	·83	2·81835	·71904	·35482	1·39074
·34	1·68130	·97448	·59478	1·02619	·84	2·85202	·71503	·35063	1·39854
0·35	1·69713	0·96783	0·58923	1·03324	0·85	2·88626	0·71106	0·34647	1·40636
·36	1·71318	·96126	·58371	1·04031	·86	2·92109	·70712	·34234	1·41419
·37	1·72946	·95475	·57821	1·04739	·87	2·95651	·70322	·33824	1·42204
·38	1·74598	·94832	·57274	1·05450	·88	2·99254	·69935	·33416	1·42989
·39	1·76273	·94196	·56730	1·06162	·89	3·02918	·69553	·33012	1·43776
0·40	1·77973	0·93567	0·56188	1·06876	0·90	3·06646	0·69173	0·32611	1·44564
·41	1·79697	·92944	·55649	1·07591	·91	3·10438	·68798	·32212	1·45354
·42	1·81447	·92329	·55113	1·08308	·92	3·14296	·68425	·31817	1·46144
·43	1·83222	·91720	·54579	1·09028	·93	3·18222	·68057	·31425	1·46936
·44	1·85023	·91118	·54047	1·09748	·94	3·22216	·67691	·31035	1·47729
0·45	1·86850	0·90522	0·53519	1·10471	0·95	3·26280	0·67329	0·30648	1·48524
·46	1·88704	·89932	·52993	1·11195	·96	3·30416	·66971	·30265	1·49319
·47	1·90586	·89349	·52470	1·11921	·97	3·34624	·66615	·29884	1·50116
·48	1·92496	·88772	·51949	1·12648	·98	3·38908	·66263	·29506	1·50914
·49	1·94434	·88201	·51431	1·13377	·99	3·43268	·65914	·29132	1·51713
0·50	1·96402	0·87636	0·50916	1·14108	1·00	3·47705	0·65568	0·28760	1·52514

$$Z(X) = e^{-\frac{1}{2}X^2}/\sqrt{(2\pi)}, \quad P(X) = 1 - Q(X) = \int_{-\infty}^{X} Z(u)\,du.$$

X	P/Z	Q/Z	Z/P	Z/Q	X	P/Z	Q/Z	Z/P	Z/Q
1·00	3·47705	0·65568	0·28760	1·52514	1·50	7·20514	0·51582	0·13879	1·93868
·01	3·52222	·65225	·28391	1·53315	·51	7·32448	·51356	·13653	1·94719
·02	3·56821	·64885	·28025	1·54118	·52	7·44636	·51133	·13429	1·95570
·03	3·61502	·64549	·27662	1·54922	·53	7·57087	·50911	·13208	1·96423
·04	3·66268	·64215	·27302	1·55726	·54	7·69805	·50690	·12990	1·97276
1·05	3·71121	0·63885	0·26945	1·56532	1·55	7·82799	0·50472	0·12775	1·98130
·06	3·76062	·63557	·26591	1·57340	·56	7·96074	·50255	·12562	1·98985
·07	3·81094	·63232	·26240	1·58148	·57	8·09639	·50040	·12351	1·99841
·08	3·86218	·62910	·25892	1·58958	·58	8·23500	·49826	·12143	2·00698
·09	3·91437	·62591	·25547	1·59768	·59	8·37664	·49614	·11938	2·01555
1·10	3·96752	0·62274	0·25205	1·60580	1·60	8·52140	0·49404	0·11735	2·02413
·11	4·02166	·61961	·24865	1·61392	·61	8·66935	·49195	·11535	2·03272
·12	4·07681	·61650	·24529	1·62206	·62	8·82058	·48988	·11337	2·04131
·13	4·13299	·61342	·24196	1·63021	·63	8·97517	·48782	·11142	2·04992
·14	4·19023	·61036	·23865	1·63837	·64	9·13320	·48578	·10949	2·05853
1·15	4·24854	0·60733	0·23538	1·64654	1·65	9·29477	0·48376	0·10759	2·06715
·16	4·30795	·60433	·23213	1·65472	·66	9·45996	·48175	·10571	2·07578
·17	4·36849	·60135	·22891	1·66292	·67	9·62887	·47975	·10385	2·08441
·18	4·43018	·59840	·22572	1·67112	·68	9·80159	·47777	·10202	2·09305
·19	4·49305	·59548	·22257	1·67933	·69	9·97824	·47580	·10022	2·10170
1·20	4·55713	0·59257	0·21944	1·68755	1·70	10·15889	0·47385	0·098436	2·11036
·21	4·62243	·58970	·21634	1·69578	·71	10·34367	·47192	·096677	2·11902
·22	4·68900	·58684	·21326	1·70403	·72	10·53268	·46999	·094943	2·12769
·23	4·75685	·58402	·21022	1·71228	·73	10·72604	·46808	·093231	2·13637
·24	4·82603	·58121	·20721	1·72054	·74	10·92384	·46619	·091543	2·14506
1·25	4·89655	0·57843	0·20422	1·72882	1·75	11·12622	0·46431	0·089878	2·15375
·26	4·96845	·57567	·20127	1·73710	·76	11·33330	·46244	·088236	2·16245
·27	5·04176	·57294	·19834	1·74539	·77	11·54520	·46058	·086616	2·17115
·28	5·11652	·57022	·19544	1·75369	·78	11·76205	·45874	·085019	2·17987
·29	5·19276	·56754	·19258	1·76201	·79	11·98397	·45692	·083445	2·18859
1·30	5·27051	0·56487	0·18974	1·77033	1·80	12·21112	0·45510	0·081893	2·19731
·31	5·34980	·56222	·18692	1·77866	·81	12·44362	·45330	·080362	2·20605
·32	5·43069	·55960	·18414	1·78700	·82	12·68163	·45151	·078854	2·21479
·33	5·51319	·55699	·18138	1·79535	·83	12·92528	·44973	·077368	2·22353
·34	5·59735	·55441	·17866	1·80371	·84	13·17474	·44797	·075903	2·23229
1·35	5·68321	0·55185	0·17596	1·81208	1·85	13·43017	0·44622	0·074459	2·24105
·36	5·77081	·54931	·17329	1·82046	·86	13·69171	·44448	·073037	2·24982
·37	5·86019	·54679	·17064	1·82884	·87	13·95955	·44275	·071636	2·25859
·38	5·95139	·54430	·16803	1·83724	·88	14·23386	·44104	·070255	2·26737
·39	6·04446	·54182	·16544	1·84564	·89	14·51481	·43934	·068895	2·27615
1·40	6·13944	0·53936	0·16288	1·85406	1·90	14·80258	0·43765	0·067556	2·28495
·41	6·23638	·53692	·16035	1·86248	·91	15·09737	·43597	·066237	2·29375
·42	6·33533	·53450	·15784	1·87091	·92	15·39937	·43430	·064938	2·30255
·43	6·43632	·53210	·15537	1·87935	·93	15·70877	·43265	·063659	2·31136
·44	6·53942	·52972	·15292	1·88780	·94	16·02580	·43100	·062399	2·32018
1·45	6·64467	0·52735	0·15050	1·89626	1·95	16·35065	0·42937	0·061160	2·32900
·46	6·75213	·52501	·14810	1·90473	·96	16·68354	·42775	·059939	2·33784
·47	6·86186	·52268	·14573	1·91320	·97	17·02472	·42614	·058738	2·34667
·48	6·97389	·52038	·14339	1·92169	·98	17·37440	·42454	·057556	2·35551
·49	7·08830	·51809	·14108	1·93018	·99	17·73282	·42295	·056393	2·36436
1·50	7·20514	0·51582	0·13879	1·93868	2·00	18·10025	0·42137	0·055248	2·37322

The normal probability function. Table of area/ordinate ratios and their reciprocals (cont.)

X	P/Z	Q/Z	Z/P	Z/Q	X	P/Z	Q/Z	Z/P	Z/Q
2·00	18·10025	0·42137	0·055248	2·37322	2·50	56·69633	0·35427	0·017638	2·82274
·01	18·47692	·41980	·054122	2·38208	·51	58·14464	·35313	·017198	2·83186
·02	18·86311	·41825	·053014	2·39094	·52	59·63565	·35199	·016768	2·84097
·03	19·25908	·41670	·051924	2·39981	·53	61·17075	·35087	·016348	2·85010
·04	19·66512	·41516	·050851	2·40869	·54	62·75138	·34975	·015936	2·85922
2·05	20·08152	0·41364	0·049797	2·41758	2·55	64·37902	0·34863	0·015533	2·86835
·06	20·50857	·41212	·048760	2·42646	·56	66·05523	·34753	·015139	2·87748
·07	20·94657	·41062	·047741	2·43536	·57	67·78159	·34643	·014753	2·88662
·08	21·39586	·40912	·046738	2·44426	·58	69·55976	·34533	·014376	2·89576
·09	21·85675	·40764	·045752	2·45317	·59	71·39146	·34425	·014007	2·90491
2·10	22·32959	0·40616	0·044784	2·46208	2·60	73·27844	0·34316	0·013647	2·91406
·11	22·81471	·40470	·043831	2·47100	·61	75·22256	·34209	·013294	2·92321
·12	23·31249	·40324	·042896	2·47992	·62	77·22570	·34102	·012949	2·93237
·13	23·82329	·40179	·041976	2·48885	·63	79·28985	·33996	·012612	2·94153
·14	24·34749	·40036	·041072	2·49778	·64	81·41704	·33890	·012282	2·95070
2·15	24·88550	0·39893	0·040184	2·50672	2·65	83·60939	0·33785	0·011960	2·95987
·16	25·43771	·39751	·039312	2·51566	·66	85·86908	·33681	·011646	2·96904
·17	26·00455	·39610	·038455	2·52461	·67	88·19839	·33577	·011338	2·97822
·18	26·58645	·39470	·037613	2·53357	·68	90·59968	·33474	·011038	2·98740
·19	27·18387	·39331	·036787	2·54253	·69	93·07536	·33371	·010744	2·99658
2·20	27·79726	0·39193	0·035975	2·55150	2·70	95·62799	0·33269	0·010457	3·00577
·21	28·42711	·39055	·035178	2·56047	·71	98·26016	·33168	·010177	3·01497
·22	29·07391	·38919	·034395	2·56944	·72	100·97461	·33067	·0²99035	3·02416
·23	29·73816	·38783	·033627	2·57842	·73	103·77414	·32967	·0²96363	3·03336
·24	30·42041	·38649	·032873	2·58741	·74	106·66167	·32867	·0²93754	3·04257
2·25	31·12118	0·38515	0·032132	2·59640	2·75	109·64022	0·32768	0·0²91207	3·05177
·26	31·84105	·38382	·031406	2·60540	·76	112·71295	·32669	·0²88721	3·06098
·27	32·58060	·38250	·030693	2·61440	·77	115·88308	·32571	·0²86294	3·07020
·28	33·34041	·38118	·029994	2·62341	·78	119·15401	·32474	·0²83925	3·07942
·29	34·12113	·37988	·029307	2·63242	·79	122·52923	·32377	·0²81613	3·08864
2·30	34·92338	0·37858	0·028634	2·64144	2·80	126·01238	0·32280	0·0²79357	3·09787
·31	35·74783	·37729	·027974	2·65046	·81	129·60721	·32184	·0²77156	3·10710
·32	36·59516	·37601	·027326	2·65948	·82	133·31763	·32089	·0²75009	3·11633
·33	37·46608	·37474	·026691	2·66851	·83	137·14770	·31994	·0²72914	3·12556
·34	38·36133	·37348	·026068	2·67755	·84	141·10162	·31900	·0²70871	3·13480
2·35	39·28165	0·37222	0·025457	2·68659	2·85	145·18375	0·31806	0·0²68878	3·14405
·36	40·22783	·37097	·024858	2·69563	·86	149·39863	·31713	·0²66935	3·15329
·37	41·20068	·36973	·024271	2·70468	·87	153·75095	·31620	·0²65040	3·16254
·38	42·20103	·36850	·023696	2·71374	·88	158·24559	·31528	·0²63193	3·17180
·39	43·22974	·36727	·023132	2·72280	·89	162·88761	·31436	·0²61392	3·18106
2·40	44·28771	0·36605	0·022580	2·73186	2·90	167·68228	0·31345	0·0²59637	3·19032
·41	45·37586	·36484	·022038	2·74093	·91	172·63504	·31254	·0²57926	3·19958
·42	46·49515	·36364	·021508	2·75000	·92	177·75156	·31164	·0²56258	3·20885
·43	47·64656	·36244	·020988	2·75908	·93	183·03773	·31074	·0²54634	3·21812
·44	48·83112	·36125	·020479	2·76816	·94	188·49965	·30985	·0²53050	3·22739
2·45	50·04988	0·36007	0·019980	2·77725	2·95	194·14366	0·30896	0·0²51508	3·23667
·46	51·30393	·35889	·019492	2·78634	·96	199·97636	·30808	·0²50006	3·24595
·47	52·59442	·35773	·019013	2·79543	·97	206·00458	·30720	·0²48543	3·25523
·48	53·92250	·35657	·018545	2·80453	·98	212·23545	·30632	·0²47118	3·26452
·49	55·28938	·35541	·018087	2·81364	·99	218·67633	·30545	·0²45730	3·27381
2·50	56·69633	0·35427	0·017638	2·82274	3·00	225·33490	0·30459	0·0²44378	3·28310

$$Z(X) = e^{-\frac{1}{2}X^2}/\sqrt{(2\pi)}, \quad P(X) = 1 - Q(X) = \int_{-\infty}^{X} Z(u)\,du.$$

X	Q/Z	Z/P	Z/Q	X	Q/Z	Z/P	Z/Q
3·00	0·30459	$0\cdot0^{3}44378$	3·28310	3·50	0·26657	$0\cdot0^{3}87289$	3·75139
·01	·30373	$\cdot0^{3}43063$	3·29240	·51	·26590	$\cdot0^{3}84281$	3·76082
·02	·30287	$\cdot0^{3}41782$	3·30169	·52	·26523	$\cdot0^{3}81370$	3·77026
·03	·30202	$\cdot0^{3}40535$	3·31100	·53	·26457	$\cdot0^{3}78551$	3·77969
·04	·30118	$\cdot0^{3}39322$	3·32030	·54	·26391	$\cdot0^{3}75822$	3·78913
3·05	0·30034	$0\cdot0^{3}38141$	3·32961	3·55	0·26326	$0\cdot0^{3}73180$	3·79857
·06	·29950	$\cdot0^{3}36992$	3·33892	·56	·26260	$\cdot0^{3}70624$	3·80802
·07	·29866	$\cdot0^{3}35874$	3·34824	·57	·26195	$\cdot0^{3}68150$	3·81746
·08	·29784	$\cdot0^{3}34787$	3·35755	·58	·26131	$\cdot0^{3}65756$	3·82691
·09	·29701	$\cdot0^{3}33729$	3·36687	·59	·26066	$\cdot0^{3}63440$	3·83636
3·10	0·29619	$0\cdot0^{3}32700$	3·37620	3·60	0·26002	$0\cdot0^{3}61200$	3·84581
·11	·29538	$\cdot0^{3}31699$	3·38552	·61	·25939	$\cdot0^{3}59033$	3·85527
·12	·29456	$\cdot0^{3}30726$	3·39485	·62	·25875	$\cdot0^{3}56936$	3·86472
·13	·29376	$\cdot0^{3}29780$	3·40418	·63	·25812	$\cdot0^{3}54909$	3·87418
·14	·29295	$\cdot0^{3}28860$	3·41352	·64	·25749	$\cdot0^{3}52949$	3·88364
3·15	0·29215	$0\cdot0^{3}27965$	3·42286	3·65	0·25686	$0\cdot0^{3}51053$	3·89311
·16	·29136	$\cdot0^{3}27096$	3·43220	·66	·25624	$\cdot0^{3}49221$	3·90257
·17	·29057	$\cdot0^{3}26251$	3·44154	·67	·25562	$\cdot0^{3}47449$	3·91204
·18	·28978	$\cdot0^{3}25430$	3·45089	·68	·25500	$\cdot0^{3}45737$	3·92151
·19	·28900	$\cdot0^{3}24632$	3·46024	·69	·25439	$\cdot0^{3}44082$	3·93098
3·20	0·28822	$0\cdot0^{3}23857$	3·46959	3·70	0·25378	$0\cdot0^{3}42482$	3·94046
·21	·28744	$\cdot0^{3}23104$	3·47895	·71	·25317	$\cdot0^{3}40937$	3·94993
·22	·28667	$\cdot0^{3}22373$	3·48830	·72	·25256	$\cdot0^{3}39444$	3·95941
·23	·28590	$\cdot0^{3}21662$	3·49767	·73	·25196	$\cdot0^{3}38002$	3·96889
·24	·28514	$\cdot0^{3}20972$	3·50703	·74	·25136	$\cdot0^{3}36608$	3·97838
3·25	0·28438	$0\cdot0^{3}20302$	3·51640	3·75	0·25076	$0\cdot0^{3}35263$	3·98786
·26	·28363	$\cdot0^{3}19652$	3·52576	·76	·25017	$\cdot0^{3}33963$	3·99735
·27	·28287	$\cdot0^{3}19020$	3·53514	·77	·24957	$\cdot0^{3}32708$	4·00683
·28	·28213	$\cdot0^{3}18407$	3·54451	·78	·24898	$\cdot0^{3}31496$	4·01632
·29	·28138	$\cdot0^{3}17812$	3·55389	·79	·24840	$\cdot0^{3}30326$	4·02582
3·30	0·28064	$0\cdot0^{3}17234$	3·56327	3·80	0·24781	$0\cdot0^{3}29197$	4·03531
·31	·27990	$\cdot0^{3}16674$	3·57265	·81	·24723	$\cdot0^{3}28107$	4·04481
·32	·27917	$\cdot0^{3}16130$	3·58203	·82	·24665	$\cdot0^{3}27054$	4·05431
·33	·27844	$\cdot0^{3}15602$	3·59142	·83	·24607	$\cdot0^{3}26039$	4·06381
·34	·27772	$\cdot0^{3}15090$	3·60081	·84	·24550	$\cdot0^{3}25059$	4·07331
3·35	0·27699	$0\cdot0^{3}14593$	3·61020	3·85	0·24493	$0\cdot0^{3}24114$	4·08281
·36	·27627	$\cdot0^{3}14112$	3·61960	·86	·24436	$\cdot0^{3}23202$	4·09232
·37	·27556	$\cdot0^{3}13644$	3·62900	·87	·24379	$\cdot0^{3}22322$	4·10183
·38	·27485	$\cdot0^{3}13191$	3·63840	·88	·24323	$\cdot0^{3}21474$	4·11134
·39	·27414	$\cdot0^{3}12752$	3·64780	·89	·24267	$\cdot0^{3}20656$	4·12085
3·40	0·27343	$0\cdot0^{3}12326$	3·65720	3·90	0·24211	$0\cdot0^{3}19866$	4·13036
·41	·27273	$\cdot0^{3}11914$	3·66661	·91	·24155	$\cdot0^{3}19106$	4·13988
·42	·27203	$\cdot0^{3}11513$	3·67602	·92	·24100	$\cdot0^{3}18372$	4·14940
·43	·27134	$\cdot0^{3}11126$	3·68544	·93	·24045	$\cdot0^{3}17665$	4·15892
·44	·27065	$\cdot0^{3}10750$	3·69485	·94	·23990	$\cdot0^{3}16983$	4·16844
3·45	0·26996	$0\cdot0^{3}10386$	3·70427	3·95	0·23935	$0\cdot0^{3}16326$	4·17796
·46	·26927	$\cdot0^{3}10033$	3·71369	·96	·23881	$\cdot0^{3}15693$	4·18749
·47	·26859	$\cdot0^{3}96911$	3·72311	·97	·23826	$\cdot0^{3}15083$	4·19702
·48	·26791	$\cdot0^{3}93601$	3·73254	·98	·23772	$\cdot0^{3}14495$	4·20654
·49	·26724	$\cdot0^{3}90394$	3·74196	·99	·23719	$\cdot0^{3}13929$	4·21608
3·50	0·26657	$0\cdot0^{3}87289$	3·75139	4·00	0·23665	$0\cdot0^{3}13383$	4·22561

REFERENCES

NATIONAL BUREAU OF STANDARDS (1953). *Tables of Normal Probability Functions*, Applied Mathematics Series, no. 23. Washington: U.S. Department of Commerce.

PEARSON, E. S. & HARTLEY, H. O. (1954). *Biometrika Tables for Statisticians*, **1**. Cambridge University Press for the Biometrika Trustees.

PEARSON, K. (1931). *Tables for Statisticians and Biometricians*, **2**. Cambridge University Press for the Biometrika Trustees.

TATE, R. F. (1955). The theory of correlation between two continuous variables when one is dichotomized. *Biometrika*, **42**. 205.

TABLES OF SYMMETRIC FUNCTIONS. PART V

By F. N. DAVID and M. G. KENDALL

1. We write M for monomial symmetric functions, S for the one-part functions, U for the unitary functions and H for the homogeneous product sums. Previously we have given tables *MS-SM* (1949), *UM-MU*, *HU-UH* (1951) and *MH-HM* (1953). We now complete the set, up to and including weight (w) 12, by the *US-SU* tables. The *HS-SH* are the same as the present tables as far as coefficients are concerned.

2. The *US-SU* tables have been partially calculated when in the course of constructing the previous sets. There are many ways in which they can be built up. We have used here both the D operator technique and the elementary relations

$$n!\,a_n = \Sigma(-1)^{n+\pi_1+\pi_2+\ldots+\pi_k}\frac{n!}{1^{\pi_1}2^{\pi_2}\ldots k^{\pi_k}\pi_1!\,\pi_2!\ldots\pi_k!}s_1^{\pi_1}s_2^{\pi_2}\ldots s_k^{\pi_k},$$

$$S_n = \Sigma(-1)^{n+\pi_1+\ldots+\pi_k}\frac{(\pi_1+\pi_2+\ldots+\pi_k-1)!\,n}{\pi_1!\,\pi_2!\ldots\pi_k!}a_1^{\pi_1}a_2^{\pi_2}\ldots a_k^{\pi_k},$$

the sum in each case being taken over all possible partitions.

3. To express the S-functions in terms of the U-functions we read horizontally up to and including the diagonal figure in bold type. Thus for $w = 5$ we have, for example,

$$(3)(2) = a_1^5 - 5a_2a_1^3 + 6a_2^2a_1 + 3a_3a_1^2 - 6a_3a_2.$$

To express the U-functions in terms of the S-functions we read vertically downwards up to and including the diagonal in bold type and divide the coefficients by $w!$. Thus

$$a_3a_2.5! = 10(1)^5 - 40(2)(1)^3 + 30(2)^2(1) + 20(3)(1)^2 - 20(3)(2).$$

4. We have called attention previously to the use of symmetric functions in distribution problems. These present tables will be useful for that purpose. They will also add considerably to the flexibility of the symmetric function system, for with this concluding set it will now be possible to express any symmetric function in terms of any other.

REFERENCES

David, F. N. & Kendall, M. G. (1949). *Biometrika*, **36**, 431.
David, F. N. & Kendall, M. G. (1951). *Biometrika*, **38**, 435.
David, F. N. & Kendall, M. G. (1953). *Biometrika*, **40**, 427.

Table 5.2

$w=2$	a_1^2	a_2
$(1)^2$	$\frac{2}{1}$	1
(2)	1	$\frac{-1}{-2}$

Table 5.3

$w=3$	a_1^3	a_1a_2	a_3
$(1)^3$	$\frac{6}{1}$	3	1
(2) (1)	1	$\frac{-3}{-2}$	−3
(3)	1	−3	$\frac{2}{3}$

Table 5.4

$w=4$	a_1^4	$a_1^2a_2$	a_2^2	a_1a_3	a_4
$(1)^4$	$\frac{24}{1}$	12	6	4	1
(2) $(1)^2$	1	$\frac{-12}{-2}$	−12	−12	−6
$(2)^2$	1	−4	$\frac{6}{4}$	.	3
(3) (1)	1	−3	.	$\frac{8}{3}$	8
(4)	1	−4	2	4	$\frac{-6}{-4}$

Table 5.5

$w=5$	a_1^5	$a_1^3a_2$	$a_1^2a_3$	$a_1a_2^2$	a_1a_4	a_2a_3	a_5
$(1)^5$	$\frac{120}{1}$	60	30	20	10	5	1
(2) $(1)^3$	1	$\frac{-60}{-2}$	−60	−60	−40	−30	−10
$(2)^2$ (1)	1	−4	$\frac{30}{4}$	.	30	15	15
(3) $(1)^2$	1	−3	.	$\frac{40}{3}$	20	40	20
(3) (2)	1	−5	6	3	$\frac{-20}{-6}$	.	−20
(4) (1)	1	−4	2	4	.	$\frac{-30}{-4}$	−30
(5)	1	−5	5	5	−5	−5	$\frac{24}{5}$

Table 5.6

$w=6$	a_1^6	$a_1^4a_2$	$a_1^3a_3$	$a_1^2a_2^2$	a_1a_4	$a_1a_2a_3$	a_2^3	a_1a_5	a_3^2	a_2a_4	a_6
$(1)^6$	$\frac{720}{1}$	360	180	90	120	60	20	30	15	6	1
(2) $(1)^4$	1	$\frac{-360}{-2}$	−360	−270	−360	−240	−120	−180	−105	−60	−15
$(2)^2$ $(1)^2$	1	−4	$\frac{180}{4}$	270	.	180	180	90	135	90	45
$(2)^3$	1	−6	12	$\frac{-90}{-8}$	.	.	.	.	−45	.	15
(3) $(1)^3$	1	−3	.	.	$\frac{240}{3}$	120	80	240	120	120	40
(3) (2) (1)	1	−5	6	.	3	$\frac{-120}{-6}$	−240	.	−120	−120	−120
$(3)^2$	1	−6	9	.	6	−18	$\frac{80}{9}$	.	.	.	40
(4) $(1)^2$	1	−4	2	.	4	.	.	$\frac{-180}{-4}$	−90	−180	−90
(4) (2)	1	−6	10	−4	4	−8	.	−4	$\frac{90}{8}$	.	90
(5) (1)	1	−5	5	.	5	−5	.	−5	.	$\frac{144}{5}$	144
(6)	1	−6	9	−2	6	−12	3	−6	6	6	$\frac{-120}{-6}$

Table 5.7

$w=7$	a_1^7	$a_1^5a_2$	$a_1^4a_3$	$a_1^3a_2^2$	$a_1^3a_4$	$a_1^2a_2a_3$	$a_1a_2^3$	$a_1^2a_5$	$a_1a_3^2$	$a_1a_2a_4$	a_1a_6	a_2a_5	a_3a_4	a_7
$(1)^7$	$\frac{5040}{1}$	2520	1260	630	840	420	210	140	210	105	35	42	21	7	1
(2) $(1)^5$	1	$\frac{-2520}{-2}$	−2520	−1890	−2520	−1680	−1050	−840	−1260	−735	−315	−420	−231	−105	−21
$(2)^2$ $(1)^3$	1	−4	$\frac{1260}{4}$	1890	.	1260	1470	1260	630	945	735	630	525	315	105
$(2)^3$ (1)	1	−6	12	$\frac{-630}{-8}$	.	.	−630	.	.	−315	−315	.	−315	−105	−105
(3) $(1)^4$	1	−3	.	.	$\frac{1680}{3}$	840	420	560	1680	840	350	840	420	280	70
(3) (2) $(1)^2$	1	−5	6	.	3	$\frac{-840}{-6}$	−840	−1680	.	.	−840	−1260	−840	−840	−840	−420
$(3)^2$ (2)	1	−7	16	−12	3	−12	$\frac{420}{12}$	.	.	.	210	.	420	.	210
$(3)^2$ (1)	1	−6	9	.	6	−18	.	$\frac{560}{9}$	.	.	560	.	.	280	280
(4) $(1)^3$	1	−4	2	.	4	.	.	.	$\frac{-1260}{-4}$	−630	−210	−1260	−630	−630	−210
(4) (2) (1)	1	−6	10	−4	4	−8	.	.	−4	$\frac{630}{8}$	630	.	630	630	630
(4) (3)	1	−7	14	−6	7	−24	6	12	−4	12	$\frac{-420}{-12}$	.	.	.	−420
(5) $(1)^2$	1	−5	5	.	5	−5	.	.	−5	.	.	$\frac{1008}{5}$	504	1008	504
(5) (2)	1	−7	15	−10	5	−15	10	.	−5	10	.	5	$\frac{-504}{-10}$	.	−504
(6) (1)	1	−6	9	−2	6	−12	.	3	−6	6	.	6	.	$\frac{-840}{-6}$	−840
(7)	1	−7	14	−7	7	−21	7	7	−7	14	−7	7	−7	−7	$\frac{720}{7}$

Table 5.8

$w=8$ (i)	a_1^8	$a_1a_1^6$	$a_2^2a_1^4$	$a_2^3a_1^2$	a_2^4	$a_3a_1^5$	$a_3a_2a_1^3$	$a_3a_2^2a_1$	$a_3^2a_1^2$	$a_3^2a_2$	$a_4a_1^4$
$(1)^8$	40320	20160	10080	5040	2520	6720	3360	1680	1120	560	1680
	1										
$(2)(1)^6$		−20160	−20160	−15120	−10080	−20160	−13440	−8400	−6720	−3920	−10080
	1	−2									
$(2)^2(1)^4$			10080	15120	15120	.	10080	11760	10080	8400	5040
	1	−4	4								
$(2)^3(1)^2$				−5040	−10080	.	.	−5040	.	−5040	.
	1	−6	12	−8							
$(2)^4$					2520	.	.	.	.	.	.
	1	−8	24	−32	16						
$(3)(1)^5$						13440	6720	3360	4480	2240	13440
	1	−3	.	.	.	3					
$(3)(2)(1)^3$							−6720	−6720	−13440	−8960	.
	1	−5	6	.	.	3	−6				
$(3)(2)^2(1)$								3360	.	6720	.
	1	−7	16	−12	.	3	−12	12			
$(3)^2(1)^2$									4480	2240	.
	1	−6	9	.	.	6	−18	.	9		
$(3)^2(2)$										−2240	.
	1	−8	21	−18	.	6	−30	36	9	−18	
$(4)(1)^4$											−10080
	1	−4	2	.	.	4	.	.	.	.	−4
$(4)(2)(1)^2$	1	−6	10	−4	.	4	−8	.	.	.	−4
$(4)(2)^2$	1	−8	22	−24	8	4	−16	16	.	.	−4
$(4)(3)(1)$	1	−7	14	−6	.	7	−24	6	12	.	−4
$(4)^2$	1	−8	20	−16	4	8	−32	16	16	.	−8
$(5)(1)^3$	1	−5	5	.	.	5	−5	.	.	.	−5
$(5)(2)(1)$	1	−7	15	−10	.	5	−15	10	.	.	−5
$(5)(3)$	1	−8	20	−15	.	8	−35	30	15	−15	−5
$(6)(1)^2$	1	−6	9	−2	.	6	−12	.	3	.	−6
$(6)(2)$	1	−8	21	−20	4	6	−24	24	3	−6	−6
$(7)(1)$	1	−7	14	−7	.	7	−21	7	7	.	−7
(8)	1	−8	20	−16	2	8	−32	24	12	−8	−8

$w=8$ (ii)	$a_4a_1a_1^3$	$a_4a_2^2$	$a_4a_2a_1$	a_4^2	$a_5a_1^3$	$a_5a_2a_1$	a_5a_3	$a_6a_1^2$	a_6a_2	a_7a_1	a_8
$(1)^8$	840	420	280	70	336	168	56	56	28	8	1
$(2)(1)^6$	−5880	−3360	−2520	−840	−3360	−1848	−728	−840	−448	−168	−28
$(2)^2(1)^4$	7560	6720	5880	2040	5040	4200	2520	2520	1680	840	210
$(2)^3(1)^2$	−2520	−5040	−2520	−2520	.	−4520	−2520	−840	−1680	−840	−420
$(2)^4$	.	1260	.	630	.	.	.	.	420	.	105
$(3)(1)^5$	6720	3360	2800	1120	6720	3360	1232	2240	1120	560	112
$(3)(2)(1)^3$	−6720	−6720	−10080	−6720	−6720	−6720	−5600	−6720	−4480	−3360	−1120
$(3)(2)^2(1)$	.	3360	1680	3360	.	3360	5040	.	3360	1680	1680
$(3)^2(1)^3$	.	.	4480	4480	.	.	2240	2240	1120	2240	1120
$(3)^2(2)$	.	.	.	.	.	.	−2240	.	−1120	.	−1120
$(4)(1)^4$	−5040	−2520	−1680	−840	−10080	−5040	−1680	−5040	−2520	−1680	−420
$(4)(2)(1)^2$	5040	5040	5040	5040	.	5040	5040	5040	5040	5040	2520
	8										
$(4)(2)^2$		−2520	.	−2520	.	.	.	.	−2520	.	−1260
	16	−16									
$(4)(3)(1)$			−3360	−6720	.	.	−3360	.	.	−3360	−3360
	12	.	−12								
$(4)^2$				2520	.	.	.	.	.	.	1260
	32	−16	−32	16							
$(5)(1)^3$					8064	4032	1344	8064	4032	4032	1344
	.	.	.	.	5						
$(5)(2)(1)$						−4032	−4032	.	−4032	−4032	−4032
	10	.	.	.	5	−10					
$(5)(3)$							2688	.	.	.	2688
	15	.	−15	.	5	−15	15				
$(6)(1)^2$								−6720	−3360	−6720	−3360
	6	.	.	.	6	.	.	−6			
$(6)(2)$									3360	.	3360
	18	−12	.	.	6	−12	.	−6	12		
$(7)(1)$										5760	5760
	14	.	−7	.	7	−7	.	−7	.	7	
(8)											−5040
	24	−8	−16	4	8	−16	8	−8	8	8	−8

Table 5.9

$w=9$ (i)	a_3^3	$a_4a_1^7$	$a_5^2a_1^5$	$a_5^3a_1^3$	$a_5^4a_1$	$a_6a_1^6$	$a_6a_3a_1^4$	$a_6a_5^2a_1^2$	$a_7a_2^3$	$a_7^2a_1^2$	$a_7^2a_3a_1$	a_8^2	$a_8a_2^2$	$a_8a_4a_1$
$(1)^9$	362880	181440	90720	45360	22680	60480	30240	15120	7560	10080	5040	1680	15120	7560
	1													
$(2)\,(1)^7$		−181440	−181440	−136080	−90720	−181440	−120960	−75600	−45360	−60480	−35280	−15120	−90720	−52920
	1	−2												
$(2)^2\,(1)^5$			90720	136080	136080	.	90720	105840	90720	90720	75600	45360	45360	68040
	1	−4	4											
$(2)^3\,(1)^3$				−45360	−90720	.	.	−45360	−75600	.	−45360	−45360	.	−22680
	1	−6	12	−8										
$(2)^4\,(1)$					22680	.	.	.	22680	.	.	.	.	.
	1	−8	24	−32	16									
$(3)\,(1)^6$						120960	60480	30240	15120	40320	20160	10080	120960	60480
	1	−3	.	.	.	3								
$(3)\,(2)\,(1)^4$							−60480	−60480	−45360	−120960	−80640	−60480	.	−60480
	1	−5	6	.	.	3	−6							
$(3)\,(2)^2\,(1)^2$								30240	45360	.	60480	90720	.	.
	1	−7	16	−12	.	3	−12	12						
$(3)\,(2)^3$									−15120	.	.	.	.	.
	1	−9	30	−44	24	3	−18	36	−24					
$(3)^2\,(1)^3$										40320	20160	20160	.	.
	1	−6	9	.	.	6	−18	.	.	9				
$(3)^2\,(2)\,(1)$											−20160	−60480	.	.
	1	−8	21	−18	.	6	−30	36	.	9	−18			
$(3)^3$												13440	.	.
	1	−9	27	−27	.	9	−54	81	.	27	−81	27		
$(4)\,(1)^5$													−90720	−45360
	1	−4	2	.	.	4	.	.	.	.	.	.	−4	
$(4)\,(2)\,(1)^3$														45360
	1	−6	10	−4	.	4	−8	.	.	.	.	.	−4	8
$(4)\,(2)^2\,(1)$														
	1	−8	22	−24	8	4	−16	16	.	.	.	.	−4	16
$(4)\,(3)\,(1)^2$	1	−7	14	−6	.	7	−24	6	.	12	.	.	−4	12
$(4)\,(3)\,(2)$	1	−9	28	−34	12	7	−38	54	−12	12	−24	.	−4	20
$(4)^2\,(1)$	1	−8	20	−16	4	8	−32	16	.	16	.	.	−8	32
$(5)\,(1)^4$	1	−5	5	.	.	5	−5	.	.	.	.	.	−5	.
$(5)\,(2)\,(1)^2$	1	−7	15	−10	.	5	−15	10	.	.	.	.	−5	10
$(5)\,(2)^2$	1	−9	29	−40	20	5	−25	40	−20	.	.	.	−5	20
$(5)\,(3)\,(1)$	1	−8	20	−15	.	8	−35	30	.	15	−15	.	−5	15
$(5)\,(4)$	1	−9	27	−30	10	9	−45	50	−10	20	−20	.	−9	40
$(6)\,(1)^3$	1	−6	9	−2	.	6	−12	.	.	3	.	.	−6	6
$(6)\,(2)\,(1)$	1	−8	21	−20	4	6	−24	24	.	3	−6	.	−6	18
$(6)\,(3)$	1	−9	27	−29	6	9	−48	63	−6	21	−45	9	−6	24
$(7)\,(1)^2$	1	−7	14	−7	.	7	−21	7	.	7	.	.	−7	14
$(7)\,(2)$	1	−9	28	−35	14	7	−35	49	−14	7	−14	.	−7	28
$(8)\,(1)$	1	−8	20	−16	2	8	−32	24	.	12	−8	.	−8	24
(9)	1	−9	27	−30	9	9	−45	54	−9	18	−27	3	−9	36

$w=9$ (ii)	$a_4a_1a_1^3$	$a_4a_4a_1$	$a_5^2a_1$	$a_6a_1^4$	$a_6a_3a_1^2$	$a_6a_3^2$	$a_6a_4a_1$	a_6a_5	$a_7a_1^2$	a_7a_2	a_8a_1
$(1)^9$	2520	1260	630	3024	1512	756	504	126	504	252	84	72	36	9
$(2)\,(1)^7$	−22680	−12600	−7560	−30240	−16032	−9072	−6552	−2016	−7560	−4032	−1512	−1512	−792	−252
$(2)^2\,(1)^5$	52920	37800	20460	45360	37800	27216	22680	9828	22680	15120	7560	7560	4536	1890
$(2)^3\,(1)^3$	−22680	−37800	−22680	.	−22680	−30240	−22680	−15120	−7560	−15120	−12600	−7560	−7560	−3780
$(2)^4\,(1)$	.	11340	5670	.	.	11340	.	5670	.	3780	3780	.	3780	945
$(3)\,(1)^6$	25200	12600	10080	60480	30240	15120	11088	3528	20160	10080	3528	5040	2520	1008
$(3)\,(2)\,(1)^4$	−90720	−57960	−60480	−60480	−60480	−45360	−50400	−27720	−60480	−40320	−22680	−30240	−17640	−10080
$(3)\,(2)^2\,(1)^2$	15120	52920	30240	.	30240	45360	45360	37800	.	30240	37800	15120	22680	15120
$(3)\,(2)^3$	.	−7560	.	.	.	−15120	.	−7560	.	.	−2520	.	−7560	.
$(3)^2\,(1)^3$	40320	20160	40320	.	.	.	20160	20160	20160	10080	10080	20160	10080	10080
$(3)^2\,(2)\,(1)$	.	−20160	.	.	.	.	−20160	−20160	.	−10080	−30240	.	−10080	−10080
$(3)^3$	.	.	.	.	.	.	.	.	.	.	6720	.	.	.
$(4)\,(1)^5$	−15120	−7560	−7560	−90720	−45360	−22680	−15120	−4536	−45360	−22680	−7560	−15120	−7560	−3780
$(4)\,(2)\,(1)^3$	45360	30240	45360	.	45360	45360	45360	30240	45360	45360	30240	45360	30240	22680
$(4)\,(2)^2\,(1)$	.	−22680	−22680	.	.	−22680	.	−22680	.	−22680	−22680	.	−22680	−11340
$(4)\,(3)\,(1)^3$	−30240	−15120	−60480	.	.	.	−30240	−45360	.	.	−15120	−30240	−15120	−30260
	−12													
$(4)\,(3)\,(2)$	−12	15120	.	.	.	.	.	15120	.	.	15120	.	15120	.
		24												
$(4)^2\,(1)$	−32	.	22680	.	.	.	.	22680	.	.	.	.	.	11340
			16											
$(5)\,(1)^4$				72576	36288	18144	12096	3024	72576	36288	12096	36288	18144	12096
		.	.	5										
$(5)\,(2)\,(1)^2$		.	.	.	−36288	−36288	−36288	−18144	.	−36288	−36288	−36288	−36288	−36288
					−10									
$(5)\,(2)^2$				5		18144	.	9072	.	.	.	.	18144	.
		.	.	5	−20	20								
$(5)\,(3)\,(1)$							24192	24192	.	.	24192	.	.	24192
	−15	.	.	5	−15	.	15							
$(5)\,(4)$								−18144	.	.	.	.	.	.
	−40	20	20	5	−20	10	20	−20						
$(6)\,(1)^3$									−60480	−30240	−10080	−60480	−30240	−30240
	.	.	.	6	.	.	.	.	−6					
$(6)\,(2)\,(1)$										30240	30240	.	30240	30240
	.	.	.	6	−12	.	.	.	−6	12				
$(6)\,(3)$											−20160	.	.	.
	−18	18	.	6	−18	.	18	.	−6	18	−18			
$(7)\,(1)^2$												51840	25920	51840
	−7	.	.	7	−7	.	.	.	−7	.	.	7		
$(7)\,(2)$													−25920	.
	−7	14	.	7	−21	14	.	.	−7	14	.	7	−14	
$(8)\,(1)$														−45360
	−16	.	4	8	−16	.	8	.	−8	8	.	8	.	−8
(9)	−27	18	9	9	−27	9	18	−9	−9	18	−9	9	−9	−9

Table 5.10

w = 10 (i)	a_1^{10}	$a_2a_1^8$	$a_3^2a_1^4$	$a_2^2a_1^6$	$a_3^4a_1^2$	a_2^5	$a_4a_1^7$	$a_4a_2a_1^5$	$a_4a_3^2a_1^3$	$a_4a_2^3a_1$
(1)^{10}	3628800	1814400	907200	453600	226800	113400	604800	302400	151200	75600
	1									
(2) (1)^8		−1814400	−1814400	−1360800	−907200	−567000	−1814400	−1209600	−756000	−453600
	1	−2								
(2)^2 (1)^6		907200	1360800	1360800	1134000	.	907200	1058400	907200	
	1	−4	4							
(2)^3 (1)^4			−453600	−907200	−1134000	.	.	−453600	−756000	
	1	−6	12	−8						
(2)^4 (1)^2				226800	567000	.	.	.	226800	
	1	−8	24	−32	16					
(2)^5					−113400	.	.	.	.	
	1	−10	40	−80	80	−32				
(3) (1)^7							1209600	604800	302400	151200
	1	−3	.	.	.	.	3			
(3) (2) (1)^5								−604800	−604800	−453600
	1	−5	6	.	.	.	3	−6		
(3) (2)^2 (1)^3								302400	453600	
	1	−7	16	−12	.	.	3	−12	12	
(3) (2)^3 (1)									−151200	
	1	−9	30	−44	24	.	3	−18	36	−24
(3)^2 (1)^4	1	−6	9	.	.	.	6	−18	.	.
(3)^2 (2) (1)^3	1	−8	21	−18	.	.	6	−30	36	.
(3)^2 (2)^2	1	−10	37	−60	36	.	6	−42	96	−72
(3)^3 (1)	1	−9	27	−27	.	.	9	−54	81	.
(4) (1)^6	1	−4	2	.	.	.	4	.	.	.
(4) (2) (1)^4	1	−6	10	−4	.	.	4	−8	.	.
(4) (2)^2 (1)^2	1	−8	22	−24	8	.	4	−16	16	.
(4) (2)^3	1	−10	38	−68	56	−16	4	−24	48	−32
(4) (3) (1)^3	1	−7	13	−6	.	.	7	−24	6	.
(4) (3) (2) (1)	1	−9	28	−34	12	.	7	−38	54	−12
(4) (3)^2	1	−10	35	−48	18	.	10	−66	120	−36
(4)^2 (1)^2	1	−8	20	−16	4	.	8	−32	16	.
(4)^2 (2)	1	−10	36	−56	36	−8	8	−48	80	−32
(5) (1)^5	1	−5	5	.	.	.	5	−5	.	.
(5) (2) (1)^3	1	−7	15	−10	.	.	5	−15	10	.
(5) (2)^2 (1)	1	−9	29	−40	20	.	5	−25	40	−20
(5) (3) (1)^2	1	−8	20	−15	.	.	8	−35	30	.
(5) (3) (2)	1	−10	36	−55	30	.	8	−51	100	−60
(5) (4) (1)	1	−9	27	−30	10	.	9	−45	50	−10
(5)^2	1	−10	35	−50	25	.	10	−60	100	−50
(6) (1)^4	1	−6	9	−2	.	.	6	−12	.	.
(6) (2) (1)^2	1	−8	21	−20	4	.	6	−24	24	.
(6) (2)^2	1	−10	37	−62	44	−8	6	−36	72	−48
(6) (3) (1)	1	−9	27	−29	6	.	9	−48	63	−6
(6) (4)	1	−10	35	−50	26	−4	10	−60	90	−32
(7) (1)^3	1	−7	14	−7	.	.	7	−21	7	.
(7) (2) (1)	1	−9	28	−35	14	.	7	−35	49	−14
(7) (3)	1	−10	35	−49	21	.	10	−63	112	−42
(8) (1)^2	1	−8	20	−16	2	.	8	−32	24	.
(8) (2)	1	−10	36	−56	34	−4	8	−48	88	−48
(9) (1)	1	−9	27	−30	9	.	9	−45	54	−9
(10)	1	−10	35	−50	25	−2	10	−60	100	−40

Table 5.10 (*cont.*)

$w=10$ (ii)	$a_3^2a_1^4$	$a_3^2a_2a_1^2$	$a_3^2a_2^2$	$a_3^2a_1$	$a_4a_1^6$	$a_4a_2a_1^4$	$a_4a_2^2a_1^2$	$a_4a_2^3$	$a_4a_3a_1^3$	$a_4a_3a_2a_1$
(1)^{10}	100800	50400	25200	16800	151200	75600	37800	18900	25200	12600
(2) (1)^8	−604800	−352800	−201600	−151200	−907200	−529200	−302400	−170100	−226800	−126000
(2)^2 (1)^6	907200	756000	554400	453600	453600	680400	604800	453600	529200	378000
(2)^3 (1)^4	.	−453600	−604800	−453600	.	−226800	−453600	−529200	−226800	−378000
(2)^4 (1)^2	.	.	226800	.	.	.	113400	283500	.	113400
(2)^5	.	.	.	.	.	.	.	−56700	.	.
(3) (1)^7	403200	201600	100800	100800	1209600	604800	302400	151200	252000	126000
(3) (2) (1)^5	−1209600	−806400	−504000	−604800	.	−604800	−604800	−453600	−907200	−579600
(3) (2)^2 (1)^3	.	604800	705600	907200	.	.	302400	453600	151200	529200
(3) (2)^3 (1)	.	.	−302400	.	.	.	.	−151200	.	−75600
(3)^2 (1)^4	403200	201600	100800	201600	.	.	.	.	403200	201600
	9									
(3)^2 (2) (1)^2		−201600	−201600	−604800	.	.	.	.	.	−201600
	9	−18								
(3)^2 (2)^2			100800	.	.	.	.	.	.	.
	9	−36	36							
(3)^2 (1)				134400	.	.	.	.	.	.
	27	−81	.	27						
(4) (1)^6					−907200	−453600	−226800	−113400	−151200	−75600
	.	.	.	.	−4					
(4) (2) (1)^4						453600	453600	340200	453600	302400
	.	.	.	.	−4	8				
(4) (2)^2 (1)^2							−226800	−340200	.	−226800
	.	.	.	.	−4	16	−16			
(4) (2)^3								113400	.	.
	.	.	.	.	−4	24	−48	32		
(4) (3) (1)^3									−302400	−151200
	12	.	.	.	−4	12	.	.	−12	
(4) (3) (2) (1)										151200
	12	−24	.	.	−4	20	−24	.	−12	24
(4) (3)^2	33	−108	18	36	−4	24	−36	.	−24	72
(4)^2 (2)^2	16	.	.	.	−8	32	−16	.	−32	.
(4)^2 (2)	16	−32	.	.	−8	48	−80	32	−32	64
(5) (1)^5	.	.	.	.	−5	.	.	.	.	.
(5) (2) (1)^3	.	.	.	.	−5	10	.	.	.	.
(5) (2)^2 (1)	.	.	.	.	−5	20	−20	.	.	.
(5) (3) (1)^2	15	−15	.	.	−5	15	.	.	−15	.
(5) (3) (2)	15	−45	30	.	−5	25	−30	.	−15	30
(5) (4) (1)	20	−20	.	.	−9	40	−30	.	−40	20
(5)^2	25	−50	25	.	−10	50	−50	.	−50	50
(6) (1)^4	3	.	.	.	−6	6	.	.	.	.
(6) (2) (1)^2	3	−6	.	.	−6	18	−12	.	.	.
(6) (2)^2	3	−12	12	.	−6	30	−48	24	.	.
(6) (3) (1)	21	−45	.	9	−6	24	−18	.	−18	18
(6) (4)	27	−60	6	12	−10	54	−72	20	−48	72
(7) (1)^3	7	.	.	.	−7	14	.	.	−7	.
(7) (2) (1)	7	−14	.	.	−7	28	−28	.	−7	14
(7) (3)	28	−84	21	21	−7	35	−42	.	−28	63
(8) (1)^2	12	−8	.	.	−8	24	−8	.	−16	.
(8) (2)	12	−32	16	.	−8	40	−56	16	−16	32
(9) (1)	18	−27	.	3	−9	36	−27	.	−27	18
(10)	25	−60	15	10	−10	50	−60	10	−40	60

Table 5.10 (*cont.*)

$n = 10$ (iii)	$a_1a_2^2$	$a_2^2a_1^2$	$a_2^3a_3$	$a_1a_3^2$	$a_1a_2a_3^2$	$a_2a_3^2a_1$	$a_1a_2a_1^3$	$a_1a_3a_2$	$a_4a_2a_1$	a_3^2	$a_4a_1^3$
(1)^{10}	4200	6300	3150	30240	15120	7560	5040	2520	1260	252	5040
(2) (1)^8	−50400	−75600	−40950	−302400	−166320	−90720	−65520	−35280	−20160	−5040	−75600
(2)^2 (1)^6	201600	264600	170100	453600	378000	272160	226800	146160	98280	32760	226800
(2)^3 (1)^4	−302400	−226800	−245700	.	−226800	−302400	−226800	−226800	−151200	−75600	−75600
(2)^4 (1)^2	113400	56700	141750	.	.	113400	.	113400	56700	56700	.
(2)^5	.	.	−28350	.	.	.	.	.	.	.	.
(3) (1)^7	50400	100800	50400	604800	302400	151200	110880	55440	35280	10080	201600
(3) (2) (1)^5	−352800	−604800	−352800	−604800	−604800	−453600	−504000	−307440	−277200	−110880	−604800
(3) (2)^2 (1)^3	655200	302400	453600	.	302400	453600	453600	478800	378000	252000	.
(3) (2)^3 (1)	−151200	.	−151200	.	.	−151200	.	−226800	−75600	−151200	.
(3)^2 (1)^4	151200	403200	201600	.	.	.	201600	100800	201600	100800	201600
(3)^2 (2) (1)^2	−504000	.	−201600	.	.	.	−201600	−201600	−201600	−201600	.
(3)^2 (2)^2	50400	.	.	.	.	.	.	100800	.	100800	.
(3)^3 (1)	.	.	.	.	.	.	.	.	.	.	.
(4) (1)^6	−25200	−75600	−37800	−907200	−453600	−226800	−151200	−75600	−45360	−15120	−453600
(4) (2) (1)^4	151200	453600	264600	.	453600	453600	453600	302400	302400	151200	453600
(4) (2)^2 (1)^2	−226800	−226800	−340200	.	.	−226800	.	−226800	−226800	−226800	.
(4) (2)^3	.	.	113400	.	.	.	.	.	.	.	.
(4) (3) (1)^3	−100800	−604800	−302400	.	.	.	−302400	−151200	−453600	−302400	.
(4) (3) (2) (1)	302400	.	302400	.	.	.	.	151200	151200	302400	.
(4) (3)^2	−100800	.	.	.	.	.	.	.	.	.	.
	−36										
(4)^2 (1)^2		226800	113400	.	.	.	.	.	226800	226800	.
	.	16									
(4)^2 (2)		.	−113400	.	.	.	.	.	.	.	.
	.	16	−32								
(5) (1)^5				725760	362880	181440	120960	60480	30240	12096	725760
	.	.	.	5							
(5) (2) (1)^3					−362880	−362880	−362880	−241920	−181440	−120960	.
	.	.	.	5	−10						
(5) (2)^2 (1)						181440	.	181440	90720	181440	.
	.	.	.	5	−20	20					
(5) (3) (1)^2							241920	120960	241920	241920	.
	.	.	.	5	−15	.	15				
(5) (3) (2)								−120960	.	−241920	.
	.	.	.	5	−25	30	15	−30			
(5) (4) (1)									−181440	−362880	.
	.	20	.	5	−20	10	20	.	−20		
(5)^2										145152	.
	.	25	.	10	−50	50	50	−50	−50	25	
(6) (1)^4											−604800
	.	.	.	6	.	.	.	.	.	.	−6
(6) (2) (1)^2	.	.	.	6	−12	.	.	.	.	.	−6
(6) (2)^2	.	.	.	6	−24	24	.	.	.	.	−6
(6) (3) (1)	.	.	.	6	−18	.	18	.	.	.	−6
(6) (4)	−12	24	−24	6	−24	12	24	.	−24	.	−6
(7) (1)^3	.	.	.	7	−7	.	.	.	.	.	−7
(7) (2) (1)	.	.	.	7	−21	14	.	.	.	.	−7
(7) (3)	−21	.	.	7	−28	21	21	−21	.	.	−7
(8) (1)^2	.	4	.	8	−16	.	8	.	.	.	−8
(8) (2)	.	4	−8	8	−32	32	8	−16	.	.	−8
(9) (1)	.	9	.	9	−27	9	18	.	−9	.	−9
(10)	−10	15	−10	10	−40	30	30	−20	−20	5	−10

Table 5.10 (*cont.*)

$w=10$ (iv)	$a_4a_3a_1^3$	$a_4a_2^3$	$a_4a_2a_1$	a_4a_4	$a_7a_1^3$	$a_7a_2a_1$	a_7a_3	$a_5a_1^2$	a_5a_2	a_5a_1	a_{10}
$(1)^{10}$	2520	1260	840	210	720	360	120	90	45	10	1
$(2)\,(1)^8$	−40320	−21420	−15120	−4410	−15120	−7920	−2880	−2520	−1305	−360	−45
$(2)^2\,(1)^6$	151200	95760	75600	28980	75600	45360	20160	18900	10710	3780	630
$(2)^3\,(1)^4$	−151200	−151200	−126000	−69300	−75600	−75600	−50400	−37800	−28350	−12600	−3150
$(2)^4\,(1)^2$	37800	94500	37800	47250	.	37800	37800	9450	23625	9450	4725
$(2)^5$	.	−18900	.	−9450	.	.	.	.	−4725	.	−945
$(3)\,(1)^7$	100800	50400	35280	10080	50400	25200	8640	10080	5040	1680	240
$(3)\,(2)\,(1)^5$	−403200	−252000	−226800	−100800	−302400	−176400	−80640	−100800	−55440	−25200	−5040
$(3)\,(2)^2\,(1)^3$	302400	352800	378000	252000	151200	226800	201600	151200	126000	75600	25200
$(3)\,(2)^3\,(1)$	.	−151200	−25200	−100800	.	−75600	−100800	.	−75600	−25200	−25200
$(3)^2\,(1)^4$	100800	50400	100800	75600	201600	100800	50400	100800	50400	33600	8400
$(3)^2\,(2)\,(1)^2$	−100800	−100800	−302400	−252000	.	−100800	−201600	−100800	−100800	−100800	−50400
$(3)^2\,(2)^2$	.	50400	.	25200	.	.	50400	.	50400	.	25200
$(3)^3\,(1)$	.	.	67200	67200	.	.	67200	.	.	22400	22400
$(4)\,(1)^6$	−226800	−113400	−75600	−20160	−151200	−75600	−25200	−37800	−18900	−7560	−1260
$(4)\,(2)\,(1)^4$	453600	340200	302400	151200	453600	302400	151200	226800	132300	75600	18900
$(4)\,(2)^2\,(1)^2$	−226800	−340200	−226800	−226800	.	−226800	−226800	−113400	−170100	−113400	−56700
$(4)\,(2)^3$	.	113400	.	75600	.	.	.	.	56700	.	18900
$(4)\,(3)\,(1)^3$	.	.	−151200	−201600	−302400	−151200	−100800	−302400	−151200	−151200	−50400
$(4)\,(3)\,(2)\,(1)$	.	.	151200	302400	.	151200	302400	.	151200	151200	151200
$(4)\,(3)^2$	.	.	.	−50400	.	.	−100800	.	.	.	−50400
$(4)^2\,(1)^2$	.	.	.	113400	.	.	.	113400	56700	113400	56700
$(4)^2\,(2)$	.	.	.	−113400	.	.	.	.	−56700	.	−56700
$(5)\,(1)^5$	362880	181400	120960	30240	362880	181440	60480	120960	60480	30240	6048
$(5)\,(2)\,(1)^3$	−362880	−362880	−362880	−181440	−362880	−362880	−241920	−362880	−241920	−181440	−60480
$(5)\,(2)^2\,(1)$	.	181440	.	90720	.	181440	181440	.	181440	90720	90720
$(5)\,(3)\,(1)^2$	.	.	241920	241920	.	.	120960	241920	120960	241920	120960
$(5)\,(3)\,(2)$	.	.	.	.	.	.	−120960	.	−120960	.	−120960
$(5)\,(4)\,(1)$	.	.	.	−181440	.	.	.	.	.	−181440	−181440
$(5)^2$	.	.	.	.	.	.	.	.	.	.	72576
$(6)\,(1)^4$	−302400	−151200	−100800	−25200	−604800	−302400	−100800	−302400	−151200	−100800	−25200
$(6)\,(2)\,(1)^2$	302400	302400	302400	151200	.	302400	302400	302400	302400	302400	151200
	12										
$(6)\,(2)^2$		−151200		−75600					−151200		−75600
	24	−24									
$(6)\,(3)\,(1)$			−201600	−201600			−201600			−201600	−201600
	18	.	−18								
$(6)\,(4)$				151200							151200
	24	−12	−24	24							
$(7)\,(1)^3$					518400	259200	86400	518400	259200	259200	86400
	.	.	.	.	7						
$(7)\,(2)\,(1)$						−259200	−259200	.	−259200	−259200	−259200
	14	.	.	.	7	−14					
$(7)\,(3)$							172800	.	.	.	172800
	21	.	−21	.	7	−21	21				
$(8)\,(1)^2$								−453600	−226800	−453600	−226800
	8	.	.	.	8	.	.	−8			
$(8)\,(2)$									226800	.	226800
	24	−16	.	.	8	−16	.	−8	16		
$(9)\,(1)$										403200	403200
	18	.	−9	.	9	−9	.	−9	.	9	
(10)											−362880
	30	−10	−20	10	10	−20	10	−10	10	10	−10

Table 5.11

$w-11$ (i)	a_1^{11}	$a_1^9a_3$	$a_1^7a_3^2$	$a_1^5a_3^3$	$a_1^4a_3^2$	$a_1^3a_3$	$a_1^7a_5$	$a_1^5a_3a_5$	$a_1^3a_3^2a_5$	$a_1^4a_5^2$	$a_1^2a_3^3a_5$	$a_1^2a_3^4$
$(1)^{11}$	39916800	19958400	9979200	4989600	2494800	1247400	6652800	3326400	1663200	831600	415800	
	1											
$(2)(1)^9$		−19958400	−19958400	−14968800	−9979200	−6237000	−19958400	−13305600	−8316000	−4989600	−2910600	
	1	−2										
$(2)^2(1)^7$			9979200	14968800	14968800	12474000	.	9979200	11642400	9979200	7484400	
	1	−4	4									
$(2)^3(1)^5$				−4989600	−9979200	−12474000	.	.	−4989600	−8316000	−9147600	
	1	−6	12	−8								
$(2)^4(1)^3$					2494800	6237000	.	.	.	2494800	5405400	
	1	−8	24	−32	16							
$(2)^5(1)$						−1247400	.	.	.	.	−1247400	
	1	−10	40	−80	80	−32						
$(3)(1)^8$							13305600	6652800	3326400	1663200	831600	
	1	−3	.	.	.	.	3					
$(3)(2)(1)^6$							−6652800	−6652800	−4989600	−3326400		
	1	−5	6	.	.	.	3	−6				
$(3)(2)^2(1)^4$							3326400	4989600	4989600			
	1	−7	16	−12	.	.	3	−12	12			
$(3)(2)^3(1)^2$								−1663200	−3326400			
	1	−9	30	−44	24	.	3	−18	36	−24		
$(3)(2)^4$										831600		
	1	−11	48	−104	112	−48	3	−24	72	−96	48	
$(3)^2(1)^5$	1	−6	9	.	.	.	6	−18	.	.	.	
$(3)^2(2)(1)^3$	1	−8	21	−18	.	.	6	−30	36	.	.	
$(3)^2(2)^2(1)$	1	−10	37	−60	36	.	6	−42	96	−72	.	
$(3)^3(1)^3$	1	−9	27	−27	.	.	9	−54	81	.	.	
$(3)^3(2)$	1	−11	45	−81	54	.	9	−72	189	−162	.	
$(4)(1)^7$	1	−4	2	.	.	.	4	.	.	.	.	
$(4)(2)(1)^5$	1	−6	10	−4	.	.	4	−8	.	.	.	
$(4)(2)^2(1)^3$	1	−8	22	−24	8	.	4	−16	16	.	.	
$(4)(2)^3(1)$	1	−10	38	−68	56	−16	4	−24	48	−32	.	
$(4)(3)(1)^4$	1	−7	14	−6	.	.	7	−24	6	.	.	
$(4)(3)(2)(1)^2$	1	−9	28	−34	12	.	7	−38	54	−12	.	
$(4)(3)(2)^2$	1	−11	46	−90	80	−24	7	−52	130	−120	24	
$(4)(3)^2(1)$	1	−10	35	−48	18	.	10	−66	120	−36	.	
$(4)^2(1)^3$	1	−8	20	−16	4	.	8	−32	16	.	.	
$(4)^2(2)(1)$	1	−10	36	−56	36	−8	8	−48	80	−32	.	
$(4)^2(3)$	1	−11	44	−76	52	−12	11	−80	172	−96	12	
$(5)(1)^6$	1	−5	5	.	.	.	5	−5	.	.	.	
$(5)(2)(1)^4$	1	−7	15	−10	.	.	5	−15	10	.	.	
$(5)(2)^2(1)^2$	1	−9	29	−40	20	.	5	−25	40	−20	.	
$(5)(2)^3$	1	−11	47	−98	100	−40	5	−35	90	−100	40	
$(5)(3)(1)^3$	1	−8	20	−15	.	.	8	−35	30	.	.	
$(5)(3)(2)(1)$	1	−10	36	−55	30	.	8	−51	100	−60	.	
$(5)(3)^2$	1	−11	44	−75	45	.	11	−83	195	−135	.	
$(5)(4)(1)^2$	1	−9	27	−30	10	.	9	−45	50	−10	.	
$(5)(4)(2)$	1	−11	45	−84	70	−20	9	−63	140	−110	20	
$(5)^2(1)$	1	−10	35	−50	25	.	10	−60	100	−50	.	
$(6)(1)^5$	1	−6	9	−2	.	.	6	−12	.	.	.	
$(6)(2)(1)^3$	1	−8	21	−20	4	.	6	−24	24	.	.	
$(6)(2)^2(1)$	1	−10	37	−62	44	−8	6	−36	72	−48	.	
$(6)(3)(1)^2$	1	−9	27	−29	6	.	9	−48	63	−6	.	
$(6)(3)(2)$	1	−11	45	−83	64	−12	9	−66	159	−132	12	
$(6)(4)(1)$	1	−10	35	−50	26	−4	10	−60	96	−32	.	
$(6)(5)$	1	−11	44	−77	55	−10	11	−77	165	−115	10	
$(7)(1)^4$	1	−7	14	−7	.	.	7	−21	7	.	.	
$(7)(2)(1)^2$	1	−9	28	−35	14	.	7	−35	49	−14	.	
$(7)(2)^2$	1	−11	46	−91	84	−28	7	−49	119	−112	28	
$(7)(3)(1)$	1	−10	35	−49	21	.	10	−63	112	−42	.	
$(7)(4)$	1	−11	44	−77	56	−14	11	−77	161	−98	14	
$(8)(1)^3$	1	−8	20	−16	2	.	8	−32	24	.	.	
$(8)(2)(1)$	1	−10	36	−56	34	−4	8	−48	88	−48	.	
$(8)(3)$	1	−11	44	−76	50	−6	11	−80	180	−120	6	
$(9)(1)^2$	1	−9	27	−30	9	.	9	−45	54	−9	.	
$(9)(2)$	1	−11	45	−84	69	−18	9	−63	144	−117	18	
$(10)(1)$	1	−10	35	−50	25	−2	10	−60	100	−40	.	
(11)	1	−11	44	−77	55	−11	11	−77	165	−110	11	

Table 5.11 (*cont.*)

w = 11 (ii)	$a_9^3a_2^2$	$a_9^3a_4a_1^2$	$a_9^3a_3^2a_1$	$a_9^3a_2^3$	$a_9^3a_5$	$a_8a_1^7$	$a_8a_3a_1^5$	$a_8a_2^3a_1^3$	$a_8a_2^5a_1$	$a_8a_5a_2^2$	$a_8a_6a_3a_1^3$
$(1)^{11}$	1108800	554400	277200	184800	92400	1663200	831600	415800	207900	277200	138600
$(2)(1)^9$	−6652800	−3880800	−2217600	−1663200	−924000	−9979200	−5821200	−3326400	−1871100	−2494800	−1386000
$(2)^2(1)^7$	9979200	8316000	6098400	4989600	3326400	4989600	7484400	6652800	4989600	5821200	4158000
$(2)^3(1)^5$	.	−4989600	−6652800	−4989600	−4989600	.	−2494800	−4989600	−5821200	−2494800	−4158000
$(2)^4(1)^3$	.	.	2494800	.	2494800	.	.	1247400	3118500	.	1247400
$(2)^5(1)$	.	.	.	.	.	.	.	.	−623700	.	.
$(3)(1)^8$	4435200	2217600	1108800	1108800	554400	13305600	6652800	3326400	1663200	2772000	1386000
$(3)(2)(1)^6$	−13305600	−8872400	−5544000	−6652800	−3880800	.	−6652800	−6652800	−4989600	−9979200	−6375600
$(3)(2)^2(1)^4$	.	6652800	7761600	9979200	8316000	.	.	3326400	4989600	1663200	5821200
$(3)(2)^3(1)^2$	.	.	−3326400	.	−4989600	.	.	.	−1663200	.	−831600
$(3)(2)^4$	.	.	.	.	.	.	.	.	.	.	.
$(3)^2(1)^5$	4435200	2217600	1108800	2217600	1108800	.	.	.	.	4435200	2217600
	9										
$(3)^2(2)(1)^3$		−2217600	−2217600	−6652800	−4435200	.	.	.	.	.	−2217600
	9	−18									
$(3)^2(2)^2(1)$			1108800	.	3326400	.	.	.	.	.	.
	9	−36	36								
$(3)^3(1)^3$				1478400	739200	.	.	.	.	.	.
	27	−81	.	27							
$(3)^3(2)$					−739200	.	.	.	.	.	.
	27	−135	162	27	−54						
$(4)(1)^7$						−9979200	−4989600	−2494800	−1247400	−1663200	−831600
	.	.	.	.	.	−4					
$(4)(2)(1)^5$							4989600	4989600	3742200	4989600	3326400
	.	.	.	.	.	−4	8				
$(4)(2)^2(1)^3$								−2494800	−3742200	.	−2494800
	.	.	.	.	.	−4	16	−16			
$(4)(2)^3(1)$									1247400	.	.
	.	.	.	.	.	−4	24	−48	32		
$(4)(3)(1)^4$										−3326400	−1663200
	12	.	.	.	.	−4	12	.	.	−12	
$(4)(3)(2)(1)^2$											1663200
	12	−24	.	.	.	−4	20	−24	.	−12	24
$(4)(3)(2)^2$	12	−48	48	.	.	−4	28	−64	48	−12	48
$(4)(3)^2(1)$	33	−108	18	36	.	−4	24	−36	.	−24	72
$(4)^2(1)^3$	16	.	.	.	.	−8	32	−16	.	−32	.
$(4)^2(2)(1)$	16	−32	.	.	.	−8	48	−80	32	−32	64
$(4)^2(3)$	40	−144	48	48	.	−8	56	−112	48	−56	192
$(5)(1)^6$	.	.	.	.	.	−5	.	.	.	.	.
$(5)(2)(1)^4$	.	.	.	.	.	−5	10	.	.	.	.
$(5)(2)^2(1)^2$	.	.	.	.	.	−5	20	−20	.	.	.
$(5)(2)^3$	.	.	.	.	.	−5	30	−60	40	.	.
$(5)(3)(1)^3$	15	−15	.	.	.	−5	15	.	.	−15	.
$(5)(3)(2)(1)$	15	−45	30	.	.	−5	25	−30	.	−15	30
$(5)(3)^2$	39	−105	135	45	−45	−5	30	−45	.	−30	90
$(5)(4)(1)^2$	20	−20	.	.	.	−9	40	−30	.	−40	20
$(5)(4)(2)$	20	−60	40	.	.	−9	58	−110	60	−40	120
$(5)^2(1)$	25	−50	25	.	.	−10	50	−50	.	−50	50
$(6)(1)^5$	3	.	.	.	.	−6	6	.	.	.	.
$(6)(2)(1)^3$	3	−6	.	.	.	−6	18	−12	.	.	.
$(6)(2)^2(1)$	3	−12	12	.	.	−6	30	−48	24	.	.
$(6)(3)(1)^2$	21	−45	.	9	.	−6	24	−18	.	−18	12
$(6)(3)(2)$	21	−87	90	9	−18	−6	36	−66	36	−18	54
$(6)(4)(1)$	27	−60	6	12	.	−10	54	−72	20	−48	72
$(6)(5)$	33	−105	75	15	−15	−11	66	−105	40	−60	120
$(7)(1)^4$	7	.	.	.	.	−7	14	.	.	−7	.
$(7)(2)(1)^2$	7	−14	.	.	.	−7	28	−28	.	−7	14
$(7)(2)^2$	7	−28	28	.	.	−7	42	−84	56	−7	28
$(7)(3)(1)$	28	−84	21	21	.	−7	35	−42	.	−28	63
$(7)(4)$	35	−112	42	28	.	−11	70	−126	56	−63	168
$(8)(1)^3$	12	−8	.	.	.	−8	24	−8	.	−16	.
$(8)(2)(1)$	12	−32	16	.	.	−8	40	−56	16	−16	32
$(8)(3)$	36	−140	96	36	−24	−8	48	−80	24	−40	120
$(9)(1)^2$	18	−27	.	3	.	−9	36	−27	.	−27	18
$(9)(2)$	18	−63	54	3	−6	−9	54	−99	54	−27	72
$(10)(1)$	25	−60	15	10	.	−10	50	−60	10	−40	60
(11)	33	−110	66	22	−11	−11	66	−110	44	−55	132

Table 5.11 (*cont.*)

$w=11$ (iii)	$a_1a_3a_7^2$	$a_1a_3^2a_7$	$a_3^2a_7^2$	$a_3^2a_5a_7$	$a_3^2a_9$	$a_5a_7^4$	$a_3a_5a_7^4$	$a_3a_5^2a_7^2$	$a_5a_9^2$	$a_3a_9a_7^2$	$a_1a_9a_5a_7$
$(1)^{11}$	69300	46200	69300	34650	11550	332640	166320	83160	41580	55440	27720
$(2)\,(1)^9$	−762300	−554400	−831600	−450450	−173250	−3326400	−1820520	−997920	−540540	−720720	−388080
$(2)^2\,(1)^7$	2772000	2217600	2910600	1871100	900900	4989600	4158000	2993760	1995840	2494800	1607760
$(2)^3\,(1)^5$	−4158000	−3326400	−2494800	−2702700	−1871100	.	−2494800	−3326400	−3160080	−2494800	−2494800
$(2)^4\,(1)^3$	2702700	1247400	623700	1559250	1351350	.	.	1247400	2286900	.	1247400
$(2)^5\,(1)$	−623700	.	.	−311850	−311850	.	.	.	−623700	.	.
$(3)\,(1)^8$	693000	554400	1108800	554400	207900	6652800	3326400	1663200	831600	1219680	609840
$(3)\,(2)\,(1)^6$	−3880800	−3880800	−6652800	−3880800	−1940400	−6652800	−6652800	−4989600	−3326400	−5544000	−3381840
$(3)\,(2)^2\,(1)^4$	6098400	7207200	3326400	4989600	4851000	.	3326400	4989600	3326400	4989600	5266800
$(3)\,(2)^3\,(1)^2$	−3326400	−1663200	.	−1663200	−2494800	.	.	−1663200	−3326400	.	−2494800
$(3)\,(2)^4$	415800	.	.	.	207900	.	.	.	831600	.	.
$(3)^2\,(1)^2$	1108800	1663200	4435200	2217600	1108800	.	.	.	.	2217600	1108800
$(3)^2\,(2)\,(1)^2$	−2217600	−5544000	.	−2217600	−4435200	.	.	.	.	−2217600	−2217600
$(3)^3\,(2)^2\,(1)$	1108800	554400	.	.	1108800	.	.	.	.	.	1108800
$(3)^3\,(1)^2$	.	1478400	.	.	1478400	.	.	.	.	.	.
$(3)^3\,(2)$	.	.	.	.	.	.	.	.	.	.	.
$(4)\,(1)^7$	−415800	−277200	−831600	−415800	−138600	−9979200	−4989600	−2494800	−1247400	−1663200	−831600
$(4)\,(2)\,(1)^5$	2079000	1663200	4989600	2910600	1247400	.	4989600	4989600	3742200	4989600	3326400
$(4)\,(2)^2\,(1)^3$	−2910600	−2494800	−2494800	−3742200	−2910600	.	.	−2494800	−3742200	.	−2494800
$(4)\,(2)^3\,(1)$	1247400	.	.	1247400	1247400	.	.	.	1247400	.	.
$(4)\,(3)\,(1)^4$	−831600	−1108800	−6652800	−3326400	−1386000	.	.	.	.	−3326400	−1663200
$(4)\,(3)\,(2)\,(1)^2$	1663200	3326400	.	3326400	4989600	.	.	.	.	.	1663200
$(4)\,(3)\,(2)^2$	−831600	.	.	.	−831600	.	.	.	.	.	.
	−48										
$(4)\,(3)^2\,(1)$	.	−1108800	.	.	−2217600	.	.	.	.	.	.
	.	−36									
$(4)^2\,(1)^3$	.	.	2494800	1247400	415800	.	.	.	.	.	.
	.	.	16								
$(4)^2\,(2)\,(1)$	.	.	.	−1247400	−1247400	.	.	.	.	.	.
	.	.	16	−32							
$(4)^2\,(3)$	.	.	.	.	831600	.	.	.	.	.	.
	−48	−96	16	−48	48						
$(5)\,(1)^6$	.	.	.	.	.	7983360	3991680	1995840	997920	1330560	665280
	.	.	.	.	.	5					
$(5)\,(2)\,(1)^4$	.	.	.	.	.	.	−3991680	−3991680	−2993760	−3991680	−2661120
	.	.	.	.	.	5	−10				
$(5)\,(2)^2\,(1)^2$	.	.	.	.	.	.	1995840	2993760	.	1995840	
	.	.	.	.	.	5	−20	20			
$(5)\,(2)^3$	.	.	.	.	.	.	.	.	−997920	.	.
	.	.	.	.	.	5	−30	60	−40		
$(5)\,(3)\,(1)^3$	.	.	.	.	.	.	.	.	2661120	1330560	
	.	.	.	.	.	5	−15	.	15		
$(5)\,(3)\,(2)\,(1)$	.	.	.	.	.	.	.	.	.	−1330560	
	.	.	.	.	.	5	−25	30	.	15	−30
$(5)^2\,(3)^2$	.	−45	.	.	.	5	−30	45	.	30	−90
$(5)\,(4)\,(1)^2$	.	.	.	.	.	5	−20	10	.	20	.
$(5)\,(4)\,(2)$	−40	.	20	−40	.	5	−30	50	−20	20	−40
$(5)^2\,(1)$	.	.	25	.	.	10	−50	50	.	50	−50
$(6)\,(1)^5$	.	.	.	.	.	6	.	.	.	.	.
$(6)\,(2)\,(1)^3$	.	.	.	.	.	6	−12	.	.	.	.
$(6)\,(2)^2\,(1)$	.	.	.	.	.	6	−24	24	.	.	.
$(6)\,(3)\,(1)^3$	.	.	.	.	.	.	.	.	.	.	.
$(6)\,(3)\,(2)$	−36	.	.	.	.	6	−18	.	.	18	.
$(6)\,(4)\,(1)$	.	−12	24	−24	.	6	−30	36	.	18	−36
$(6)\,(5)$	−30	−15	30	−30	.	6	−24	12	.	24	.
$(7)\,(1)^4$	.	.	.	.	.	11	−60	75	−10	60	−90
$(7)\,(2)\,(1)^2$	.	.	.	.	.	7	−7	.	.	.	.
$(7)\,(2)^2$	−28	.	.	.	.	7	−21	14	.	.	.
$(7)\,(3)\,(1)$	.	−21	.	.	.	7	−35	50	−28	.	.
$(7)\,(4)$	−42	−56	28	−56	28	7	−28	21	.	21	−21
$(8)\,(1)^3$	.	.	4	.	.	8	−16	.	.	8	−28
$(8)\,(2)\,(1)$	.	.	4	−8	.	8	−32	32	.	8	−16
$(8)\,(3)$	−24	−48	4	−12	12	8	−40	48	.	32	−72
$(9)\,(1)^2$	.	.	9	.	.	9	−27	9	.	18	.
$(9)\,(2)$	−36	.	9	−18	.	9	−45	63	−18	18	−36
$(10)\,(1)$	.	−10	15	−10	.	10	−40	30	.	30	−20
(11)	−33	−33	22	−33	11	11	−55	66	−11	44	−66

Table 5.11 (*cont.*)

$w = 11$ (iv)	$a_1a_{10}^2$	$a_2a_3a_{6}^2$	$a_2a_4a_5$	$a_3^2a_5$	$a_4a_7^2$	$a_4a_6a_7^2$	$a_5a_6^2a_7$	$a_5a_8a_9^2$	$a_6a_7a_8$	$a_6a_9a_7$	a_8a_9
$(1)^{11}$	9240	13860	6930	2772	55440	27720	13860	9240	4620	2310	462
$(2)\,(1)^9$	−147840	−221760	−117810	−55440	−831600	−443520	−235620	−166320	−87780	−48510	−11550
$(2)^2\,(1)^7$	776160	1081080	651420	360360	2494800	1663200	1053360	831600	498960	318780	97020
$(2)^3\,(1)^5$	−1663200	−1663200	−1372140	−831600	−831600	−1663200	−1663200	−1386000	−1108800	−762300	−318780
$(2)^4\,(1)^3$	1247400	623700	1143450	623700	.	415800	1039500	415800	900900	519750	381150
$(2)^5\,(1)$	.	.	−311850	.	.	.	−207900	.	−207900	−103950	−103950
$(3)\,(1)^8$	221760	388080	194040	110880	2217600	1108800	554400	388080	194040	110880	27720
$(3)\,(2)\,(1)^6$	−1774080	−3049200	−1718640	−1219680	−6652800	−4435200	−2773000	−2494800	−1441440	−1108800	−388080
$(3)\,(2)^2\,(1)^4$	4435200	4158000	3603600	2772000	.	3326400	3880800	4158000	3326400	2772000	1386000
$(3)\,(2)^3\,(1)^2$	−3326400	−831600	−2494800	−1663200	.	.	−1663200	−277200	−2217600	−1108800	−1386000
$(3)\,(2)^4$	.	.	415800	.	.	.	.	.	138600	.	138600
$(3)^2\,(1)^5$	776160	2217600	1108800	1108800	2217600	1108800	554400	1108800	554400	831600	388080
$(3)^2\,(2)\,(1)^3$	−3326400	−2217600	−2217600	−2217600	.	−1108800	−1108800	−3326400	−2217600	−2772000	−1663200
$(3)^2\,(2)^2\,(1)$	2772000	.	1108800	1108800	.	.	554400	.	1663200	2772000	1386000
$(3)^3\,(1)^2$	739200	.	.	.	.	.	.	739200	369600	739200	369600
$(3)^3\,(2)$	−739200	.	.	.	.	.	.	.	−369600	.	−369600
$(4)\,(1)^7$	−277200	−498960	−249480	−166320	−498960	−249480	−1247400	−831600	−415800	−221760	−55440
$(4)\,(2)\,(1)^5$	1663200	3326400	1912680	1663200	4989600	4989600	3742200	3326400	2079000	1663200	665280
$(4)\,(2)^2\,(1)^3$	−2494800	−2494800	−2910600	−2494800	.	−2494800	−3742200	−2494800	−2910600	−2494800	−1663200
$(4)\,(2)^3\,(1)$	.	.	1247400	.	.	.	1247400	.	1247400	831600	831600
$(4)\,(3)\,(1)^4$	−1108800	−4989600	−2494800	−3326400	.	.	.	−1663200	−831600	−2217600	−1386000
$(4)\,(3)\,(2)\,(1)^2$	3326400	1663200	3326400	3326400	.	.	.	1663200	1663200	3326400	3326400
$(4)\,(3)\,(2)^2$	.	.	−831600	.	.	.	.	.	−831600	.	−831600
$(4)\,(3)^2\,(1)$	−1108800	.	.	.	.	.	.	.	.	−554400	−554400
$(4)^2\,(1)^3$	.	2494800	1247400	2494800	.	.	.	.	.	1247400	1247400
$(4)^2\,(2)\,(1)$	.	.	−1247400	.	.	.	.	.	.	−1247400	−1247400
$(4)^2\,(3)$	.	.	.	.	.	.	.	.	.	.	−166320
$(5)\,(1)^6$	221760	332640	166320	133056	7983360	3991680	1995840	1330560	665280	332640	77616
$(5)\,(2)\,(1)^4$	−1330560	−1995840	−1164240	−1330560	.	−3991680	−3991680	−3991680	−2661120	−1995840	−831600
$(5)\,(2)^2\,(1)^2$	1995840	997920	1496880	1995840	.	.	1995840	.	1995840	997920	1496880
$(5)\,(2)^3$	.	.	−498960	.	.	.	.	.	.	.	−166320
$(5)\,(3)\,(1)^3$	887040	2661120	1330560	2661120	.	.	.	2661120	1330560	2661120	1774080
$(5)\,(3)\,(2)\,(1)$	−2661120	.	−1330560	−2661120	.	.	.	.	−1330560	.	−2661120
$(5)\,(3)^2$	887040	.	.	.	.	.	.	.	.	.	443520
	45										
$(5)\,(4)\,(1)^3$	.	−1995840	−997920	−3991680	.	.	.	.	.	−1995840	−2993760
	.	−20									
$(5)\,(4)\,(2)$	.	.	997920	.	.	.	.	.	.	.	997920
	.	−20	40								
$(5)^2\,(1)$	.	.	.	1596672	.	.	.	.	.	.	1596672
	.	−50	.	25							
$(6)\,(1)^5$					−6652800	−3326400	−1663200	−1108800	−554400	−277200	−55440
	.	.	.	.	−6						
$(6)\,(2)\,(1)^3$						3326400	3326400	3326400	2217600	1663200	554400
	.	.	.	.	−6	12					
$(6)\,(2)^2\,(1)$							−1663200	.	−1663200	−831600	−831600
	.	.	.	.	−6	24	−24				
$(6)\,(3)\,(1)^3$								−2217600	−1108800	−2217600	−1108800
	.	.	.	.	−6	18	.	−18			
$(6)\,(3)\,(2)$									1108800	.	1108800
	.	.	.	.	−6	30	−36	−18	36		
$(6)\,(4)\,(1)$										1663200	1663200
	.	−24	.	.	−6	24	−12	−24	.	24	
$(6)\,(5)$											−1330560
	15	−60	30	30	−6	30	−30	−30	30	30	−30
$(7)\,(1)^4$	.	.	.	.	−7	.	.	.	.	.	.
$(7)\,(2)\,(1)^2$	.	.	.	.	−7	14	.	.	.	.	.
$(7)\,(2)^2$	.	.	.	.	−7	28	−28	.	.	.	.
$(7)\,(3)\,(1)$	.	.	.	.	−7	21	.	−21	.	.	.
$(7)\,(4)$	.	−28	28	.	−7	28	−14	−28	.	28	.
$(8)\,(1)^3$	.	.	.	.	−8	8	.	.	.	.	.
$(8)\,(2)\,(1)$	.	.	.	.	−8	24	−16	.	.	.	.
$(8)\,(3)$	24	.	.	.	−8	32	−24	−24	24	.	.
$(9)\,(1)^2$	.	−9	.	.	−9	18	.	−9	.	.	.
$(9)\,(2)$	.	−9	18	.	−9	36	−36	−9	18	.	.
$(10)\,(1)$	.	−20	.	5	−10	30	−10	−20	.	10	.
(11)	11	−33	22	11	−11	44	−33	−33	22	22	−11

Table 5.11 (*cont.*)

$w-11$ (v)	$a_1a_1^4$	$a_1a_1a_2^2$	$a_1a_2^2$	$a_1a_2a_1$	a_1a_4	$a_4a_2^2$	$a_4a_4a_1$	a_4a_6	$a_4a_2^3$	a_6a_5	$a_{10}a_1$	a_{11}
$(1)^{11}$	7920	3960	1980	1320	330	990	495	165	110	55	11	1
$(2)\,(1)^9$	−166320	−87120	−45540	−31680	−8910	−27720	−14355	−5115	−3960	−2035	−495	−55
$(2)^2\,(1)^7$	831600	498960	293040	221760	77220	207900	117810	48510	41580	22770	6930	990
$(2)^3\,(1)^5$	−831600	−831600	−665280	−554400	−263340	−415800	−311850	−173250	−138600	−90090	−34650	−6930
$(2)^4\,(1)^3$	.	415800	623700	415800	311850	103950	259875	225225	103950	121275	51975	17325
$(2)^5\,(1)$	.	.	−207900	.	−103950	.	−51975	−51975	.	−51975	−10395	−10395
$(3)\,(1)^8$	554400	277200	138600	95040	25740	110880	55440	18810	18480	9240	2640	330
$(3)\,(2)\,(1)^6$	−3326400	−1940400	−1108800	−887040	−332640	−1108800	−600840	−249480	−277200	−147840	−55440	−9240
$(3)\,(2)^2\,(1)^4$	1663200	2494800	2217600	2217600	1247400	1663200	1386000	900900	831600	554400	277200	69300
$(3)\,(2)^3\,(1)^2$	.	−831600	−1663200	−1108800	−1108800	.	−831600	−970200	−277200	−554400	−277200	−138600
$(3)\,(2)^4$	.	.	415800	.	207900	.	.	34650	.	138600	.	34650
$(3)^2\,(1)^6$	2217600	1108800	554400	554400	277200	1108800	554400	221760	369600	184800	92400	18480
$(3)^2\,(2)\,(1)^4$	.	−1108800	−1108800	−2217600	−1663200	−1108800	−1108800	−1108800	−1108800	−739200	−554400	−184800
$(3)^2\,(2)^2\,(1)$	.	.	554400	554400	831600	.	554400	1108800	.	554400	277200	277200
$(3)^3\,(1)^3$	.	.	.	739200	739200	.	.	369600	246400	123200	246400	123200
$(3)^3\,(2)$	.	.	.	.	.	.	.	−369600	.	−123200	.	−123200
$(4)\,(1)^7$	−1663200	−831600	−415800	−277200	−71280	−415800	−207900	−69300	−83160	−41580	−13860	−1980
$(4)\,(2)\,(1)^5$	4989600	3326400	2079000	1663200	665280	2494800	1455300	623700	831600	457380	207900	41580
$(4)\,(2)^2\,(1)^3$	.	−2494800	−2910600	−2494800	−1663200	−1247400	−1871100	−1455300	−1247400	−1039500	−623700	−207900
$(4)\,(2)^3\,(1)$	.	.	1247400	.	831600	.	623700	623700	.	623700	207900	207900
$(4)\,(3)\,(1)^5$	−3326400	−1663200	−831600	−1108800	−831600	−3326400	−1663200	−693000	−1663200	−831600	−554400	−138600
$(4)\,(3)\,(2)\,(1)^3$	.	1663200	1663200	3326400	3326400	.	1663200	2494800	1663200	1663200	1663200	831600
$(4)\,(3)\,(2)^2\,(1)$	.	.	−831600	.	−831600	.	.	−415800	.	−831600	.	−415800
$(4)\,(3)^2\,(1)$	.	.	.	−1108800	−1663200	.	.	−1108800	.	.	−554400	−554400
$(4)^2\,(1)^3$	.	.	.	.	415800	1247400	623700	207900	1247400	623700	623700	207900
$(4)^2\,(2)\,(1)$	.	.	.	.	−1247400	.	−623700	−623700	.	−623700	−623700	−623700
$(4)^2\,(3)$	.	.	.	.	831600	.	.	415800	.	.	.	415800
$(5)\,(1)^6$	3991680	1995840	997920	665280	166320	1330560	665280	221760	332640	166320	665280	11088
$(5)\,(2)\,(1)^4$	−3991680	−3991680	−2993760	−2661120	−1164240	−3991680	−2661120	−1330560	−1995840	−1164240	−665280	−166320
$(5)\,(2)^2\,(1)^2$	.	1995840	2993760	1995840	1496880	.	1995840	1995840	997920	1496880	997920	498960
$(5)\,(2)^3$	.	.	−997920	.	−498960	.	.	.	.	−498960	.	−166320
$(5)\,(3)\,(1)^3$	.	.	.	1330560	1330560	2661120	1330560	887040	2661120	1330560	1330560	443520
$(5)\,(3)\,(2)\,(1)$	.	.	.	−1330560	−1330560	.	−1330560	−2661120	.	−1330560	−1330560	−1330560
$(5)\,(3)^2$	.	.	.	.	.	.	.	887040	.	.	.	443520
$(5)\,(4)\,(1)^3$	.	.	.	.	−997920	.	.	.	−1995840	−997920	−1995840	−997920
$(5)\,(4)\,(2)$	.	.	.	.	997920	.	.	.	.	997920	.	997920
$(5)^2\,(1)$	.	.	.	.	.	.	.	.	.	.	798336	798336
$(6)\,(1)^5$	−6652800	−3326400	−1663200	−1108800	−277200	−3326400	−1663200	−554400	−1108800	−554400	−277200	−55440
$(6)\,(2)\,(1)^3$	.	3326400	3326400	3326400	1663200	3326400	3326400	2217600	3326400	2217600	1663200	554400
$(6)\,(2)^2\,(1)$	.	.	−1663200	.	−831600	.	−1663200	−1663200	.	−1663200	−831600	−831600
$(6)\,(3)\,(1)^3$	.	.	.	2217600	−2217600	.	.	−1108800	−2217600	−1108800	−2217600	−1108800
$(6)\,(3)\,(2)$	.	.	.	.	.	.	.	1108800	.	1108800	.	1108800
$(6)\,(4)\,(1)$	.	.	.	.	1663200	.	.	.	.	.	1663200	1663200
$(6)\,(5)$	.	.	.	.	.	.	.	.	.	.	.	−1330560
$(7)\,(1)^4$	5702400	2851200	1425600	950400	237600	5702400	2851200	950400	2851200	1425600	950400	237600
	7											
$(7)\,(2)\,(1)^2$		−2851200	−2851200	−2851200	−1425600	.	−2851200	−2851200	−2851200	−2851200	−2851200	−1425600
	7	−14										
$(7)\,(2)^2$			1425600	.	712800	.	.	.	.	1425600	.	712800
	7	−28	28									
$(7)\,(3)\,(1)$				1900800	1900800	.	.	1900800	.	.	1900800	1900800
	7	−21	.	21								
$(7)\,(4)$					−1425600	.	.	.	.	.	.	−1425600
	7	−28	14	28	−28							
$(8)\,(1)^3$						−4989600	−2494800	−831600	−4989600	−2494800	−2494800	−831600
	8	.	.	.	.	−8						
$(8)\,(2)\,(1)$							2494800	2494800	.	2494800	2494800	2494800
	8	−16	.	.	.	−8	16					
$(8)\,(3)$								−1663200	.	.	.	−1663200
	8	−24	.	24	.	−8	24	−24				
$(9)\,(1)^2$									4435200	2217600	4435200	2217600
	9	−9	.	.	.	−9	.	.	9			
$(9)\,(2)$										−2217600	.	−2217600
	9	−27	18	.	.	−9	18	.	9	−18		
$(10)\,(1)$											−3991680	−3991680
	10	−20	.	10	.	−10	10	.	10	.	−10	
(11)												3628800
	11	−33	11	22	−11	−11	22	−11	11	−11	−11	11

Table 5.12

| $w=12$ (i) | a_1^{12} | $a_1^{10}a_2$ | $a_1^8a_2^2$ | $a_1^6a_2^3$ | $a_1^4a_2^4$ | $a_1^2a_2^5$ | a_2^6 | $a_1^8a_3$ | $a_1^6a_2a_3$ | $a_1^4a_2^2a_3$ | $a_1^2a_2^3a_3$ | $a_2^4a_3$ | $a_1^4a_3^2$ | $a_1^2a_2a_3^2$ | $a_2^2a_3^2$ | $a_1^2a_3^3$ | $a_2a_3^3$ | a_3^4 | $a_1^6a_4$ | $a_1^4a_2a_4$ | $a_1^2a_2^2a_4$ | $a_2^3a_4$ | $a_1^2a_3a_4$ | $a_2a_3a_4$ | $a_3^2a_4$ | $a_1^2a_4^2$ | $a_2a_4^2$ | $a_3a_4^2$ | a_4^3 | $a_1^4a_5$ | $a_1^2a_2a_5$ | $a_2^2a_5$ | $a_1^2a_3a_5$ | $a_2a_3a_5$ | $a_3^2a_5$ | $a_1^2a_4a_5$ | $a_2a_4a_5$ | $a_3a_4a_5$ | $a_4^2a_5$ | $a_1^2a_5^2$ | $a_2a_5^2$ | $a_3a_5^2$ | $a_4a_5^2$ | a_5^3 | $a_1^3a_6$ | $a_1a_2a_6$ | a_3a_6 | $a_1a_4a_6$ | a_5a_6 | $a_1^2a_7$ | a_2a_7 | a_3a_7 | a_4a_7 | a_1a_8 | a_2a_8 | a_4a_8 | a_1a_9 | a_3a_9 | a_5a_9 | a_7a_9 | a_2a_{10} | a_4a_{10} | a_6a_{10} | a_8a_{10} | a_1a_{11} | a_3a_{11} | a_5a_{11} | a_7a_{11} | a_9a_{11} | $a_{10}a_{11}$ | a_{12} |
|---|---|---

Table 5.12 (*cont.*)

$w=12$ (ii)	$a_0a_2^4a_4$	$a_0^3a_1^6$	$a_0^2a_0a_1^4$	$a_0^3a_1^3a_2^3$	$a_0^5a_2^3$	$a_0^3a_1^3$	$a_0^3a_1a_3$	a_0^4	$a_1a_3^4$	$a_1a_2a_3^4$	$a_1a_2^3a_3^4$
$(1)^{12}$	4989600	13305600	6652800	3326400	1663200	2217600	1108800	369600	19958400	9979200	4989600
$(2)\,(1)^{10}$	−34927200	−79833600	−46569600	−26611200	−14968800	−19958400	−11088000	−4435200	−119750400	−69854400	−39916800
$(2)^2\,(1)^8$	80812800	119750400	99792000	73180800	49896000	59875200	39916800	19958400	59875200	89812800	79833600
$(2)^3\,(1)^6$	−109771200	.	−59875200	−79833600	−76507200	−59875200	−59875200	−39916800	.	−29937600	−59875200
$(2)^4\,(1)^4$	64864800	.	.	29937600	54885600	.	29937600	29937600	.	.	14968800
$(2)^5\,(1)^2$	−14968800	.	.	.	−14968800	.	.	.	.	.	.
$(2)^6$	.	.	.	.	.	.	.	.	.	.	.
$(3)\,(1)^9$	9979200	53222400	26611200	13305600	6652800	13305600	6652800	2956800	159667200	79833600	39916800
$(3)\,(2)\,(1)^7$	−39916800	−159667200	−106444800	−66528000	−39916800	−79833600	−46569600	−26611200	.	−79833600	−79833600
$(3)\,(2)^2\,(1)^5$	59875200	.	79833600	93139200	79833600	119750400	99792000	79833600	.	.	39916800
$(3)\,(2)^3\,(1)^3$	−39916800	.	.	−39916800	−66528000	.	−59875200	−79833600	.	.	.
$(3)\,(2)^4\,(1)$	9979200	.	.	.	19958400	.	.	.	.	.	.
	48										
$(3)^2\,(1)^6$	.	53222400	26611200	13305600	6652800	26611200	13305600	8870400	.	.	.
	.	9									
$(3)^2\,(2)\,(1)^4$	.	.	−26611200	−26611200	−19958400	−79833600	−53222400	−53222400	.	.	.
	.	9	−18								
$(3)^2\,(2)^2\,(1)^2$	.	.	.	13305600	19958400	.	39916800	79833600	.	.	.
	.	9	−36	36							
$(3)^2\,(2)^3$	.	.	.	.	−6652800	.	.	.	.	.	.
	144	9	−54	108	−72						
$(3)^3\,(1)^3$	.	.	.	.	.	17740800	8870400	11827200	.	.	.
	.	27	−81	.	.	27					
$(3)^3\,(2)\,(1)$	.	.	.	.	.	.	−8870400	−35481600	.	.	.
	.	27	−135	162	.	27	−54				
$(3)^4$	.	.	.	.	.	.	.	5913600	.	.	.
	.	54	−324	486	.	108	−324	81			
$(4)\,(1)^8$	.	.	.	.	.	.	.	.	−119750400	−59875200	−29937600
	.	.	.	.	.	.	.	.	−4		
$(4)\,(2)\,(1)^6$	.	.	.	.	.	.	.	.	.	59875200	59875200
	.	.	.	.	.	.	.	.	−4	8	
$(4)\,(2)^2\,(1)^4$	.	.	.	.	.	.	.	.	.	.	−29937600
$(4)\,(2)^3\,(1)^2$	.	.	.	.	.	.	.	.	−4	16	−16
$(4)\,(2)^4$	64	.	.	.	.	.	.	.	−4	24	−48
$(4)\,(3)\,(1)^5$	.	12	.	.	.	.	.	.	−4	32	−96
$(4)\,(3)\,(2)\,(1)^3$	.	12	−24	.	.	.	.	.	−4	12	.
$(4)\,(3)\,(2)^2\,(1)$	24	12	−48	48	.	.	.	.	−4	20	−24
$(4)\,(3)^2\,(1)^2$	.	33	−108	18	.	36	.	.	−4	28	−64
$(4)\,(3)^2\,(2)$	72	33	−174	234	−36	36	−72	.	−4	24	−36
$(4)^2\,(1)^4$	.	16	.	.	.	.	.	.	−4	32	−84
$(4)^2\,(2)^2$	.	16	−32	.	.	.	.	.	−8	32	−16
$(4)^2\,(3)\,(1)^2$	64	16	−64	64	.	.	.	.	−8	48	−80
$(4)^2\,(3)^2$	12	40	−144	48	.	48	.	.	−8	64	−176
$(4)^3\,(1)^2$	48	48	−192	96	.	64	.	.	−8	56	−112
$(4)^3\,(2)$	.	.	.	.	.	.	.	.	−12	96	−240
$(5)\,(1)^7$	.	.	.	.	.	.	.	.	−5	.	.
$(5)\,(2)\,(1)^5$	.	.	.	.	.	.	.	.	−5	10	.
$(5)\,(2)^2\,(1)^3$	.	.	.	.	.	.	.	.	−5	20	−20
$(5)\,(2)^3\,(1)$	40	.	.	.	.	.	.	.	−5	30	−60
$(5)\,(3)\,(1)^4$	.	15	−15	.	.	.	.	.	−5	15	.
$(5)\,(3)\,(2)\,(1)^2$	.	15	−45	30	.	.	.	.	−5	25	−30
$(5)\,(3)\,(2)^2$	120	15	−75	120	−60	.	.	.	−5	35	−80
$(5)\,(3)^2\,(1)$	.	39	−105	135	.	45	−45	.	−5	30	−45
$(5)\,(4)\,(1)^3$	.	20	−20	.	.	.	.	.	−9	40	−30
$(5)\,(4)\,(2)\,(1)$	20	20	−60	40	.	.	.	.	−9	58	−110
$(5)\,(4)\,(3)$	60	47	−215	210	−30	60	−60	.	−9	67	−150
$(5)^2\,(1)^2$	.	25	−50	25	.	.	.	.	−10	50	−50
$(5)^2\,(2)$	100	25	−100	125	−50	.	.	.	−10	70	−150
$(6)\,(1)^6$	.	3	.	.	.	.	.	.	−6	6	.
$(6)\,(2)\,(1)^4$	.	3	−6	.	.	.	.	.	−6	18	−12
$(6)\,(2)^2\,(1)^2$	.	3	−12	12	.	.	.	.	−6	30	−48
$(6)\,(2)^3$	96	3	−18	36	−24	.	.	.	−6	42	−108
$(6)\,(3)\,(1)^3$	.	21	−45	.	.	9	.	.	−6	24	−18
$(6)\,(3)\,(2)\,(1)$	12	21	−87	90	.	9	−18	.	−6	36	−66
$(6)\,(3)^2$	36	48	−252	324	−18	72	−162	27	−6	42	−90
$(6)\,(4)\,(1)^2$	.	27	−60	6	.	12	.	.	−10	54	−72
$(6)\,(4)\,(2)$	64	27	−114	126	−12	12	−24	.	−10	74	−180
$(6)\,(5)\,(1)$	10	33	−105	75	.	15	−15	.	−11	66	−105
$(6)^2$	48	42	−180	198	−12	30	−72	9	−12	84	−180
$(7)\,(1)^5$	.	7	.	.	.	.	.	.	−7	14	.
$(7)\,(2)\,(1)^3$	.	7	−14	.	.	.	.	.	−7	28	−28
$(7)\,(2)^2\,(1)$	28	7	−28	28	.	.	.	.	−7	42	−84
$(7)\,(3)\,(1)^2$	.	28	−84	21	.	21	.	.	−7	35	−42
$(7)\,(3)\,(2)$	84	28	−140	189	−42	21	−42	.	−7	49	−112
$(7)\,(4)\,(1)$	14	35	−112	42	.	28	.	.	−11	70	−126
$(7)\,(5)$	70	42	−175	175	−35	35	−35	.	−12	84	−175
$(8)\,(1)^4$	.	12	−8	.	.	.	.	.	−8	24	−8
$(8)\,(2)\,(1)^2$	.	12	−32	16	.	.	.	.	−8	40	−56
$(8)\,(2)^2$	96	12	−56	80	−32	.	.	.	−8	56	−136
$(8)\,(3)\,(1)$	6	36	−140	96	.	36	−24	.	−8	48	−80
$(8)\,(4)$	56	44	−184	152	−16	48	−32	.	−12	88	−200
$(9)\,(1)^3$	.	18	−27	.	.	3	.	.	−9	36	−27
$(9)\,(2)\,(1)$	18	18	−63	54	.	3	−6	.	−9	54	−99
$(9)\,(3)$	54	45	−216	243	−27	57	−90	9	−9	63	−135
$(10)\,(1)^2$	.	25	−60	15	.	10	.	.	−10	50	−60
$(10)\,(2)$	80	25	−110	135	−30	10	−20	.	−10	70	−160
$(11)\,(1)$	11	33	−110	66	.	22	−11	.	−11	66	−110
(12)	60	42	−180	180	−24	40	−48	3	−12	84	−180

Table 5.12 (*cont.*)

$w = 12$ (iii)	$a_1a_2^3a_3^3$	$a_1a_2^4$	$a_1a_2a_3^5$	$a_1a_2a_3a_4^3$	$a_1a_2a_3^3a_4$	$a_1a_2^3a_3^2$	$a_1a_2^2a_3^3$	$a_1a_2^3a_4$	$a_1^2a_3^4$	$a_1^2a_2a_3^3$	$a_1^3a_2^2$	$a_1^3a_2a_4$
$(1)^{12}$	2494800	1247400	3326400	1663200	831600	554400	277200	831600	415800	207900	138600	
$(2)\,(1)^{10}$	−22453200	−12474000	−29937600	−16632000	−9147600	−6652800	−3603600	−9979200	−5405400	−2910600	−2079000	
$(2)^2\,(1)^8$	59875200	41164200	69854400	49896000	33264000	26611200	16632000	34927200	22453200	13920300	10810800	
$(2)^3\,(1)^6$	−69854400	−64864800	−29937600	−49896000	−49896000	−39916800	−33264000	−29937600	−32432400	−27442800	−22453200	
$(2)^4\,(1)^4$	37422000	53638200	.	14968800	32432400	14968800	27442800	7484400	18711000	25571700	16216200	
$(2)^5\,(1)^2$	−7484400	−22453200	.	.	−7484400	.	−7484400	.	−3742200	−11226600	−3742200	
$(2)^6$	.	3742200	.	.	.	.	.	.	.	1871100	.	
$(3)\,(1)^9$	19958400	9979200	33264000	16632000	8316000	6652800	3326400	13305600	6652800	3326400	2494800	
$(3)\,(2)\,(1)^7$	−59875200	−39916800	−119750400	−76507200	−46509600	−46509600	−26611200	−79833600	−46509600	−26611200	−23284800	
$(3)\,(2)^2\,(1)^5$	59875200	59875200	19958400	69854400	73180800	86486400	66528000	39916800	59875200	53222400	58212000	
$(3)\,(2)^3\,(1)^3$	−19958400	−39916800	.	−9979200	−39916800	−19958400	−53222400	.	−19958400	−39916800	−29937600	
$(3)\,(2)^4\,(1)$	.	9979200	.	.	4989600	.	9979200	.	.	9979200	2494800	
$(3)^2\,(2)\,(1)^4$	.	.	53222400	26611200	13305600	19958400	9979200	53222400	26611200	13305600	13305600	
$(3)^2\,(2)^2\,(1)^2$	.	.	.	−26611200	−26611200	−66528000	−43243200	.	−26611200	−26611200	−53224400	
$(3)^2\,(2)^3$	.	.	.	.	13305600	6652800	36590400	.	.	13305600	13305600	
$(3)^3\,(2)^2$	.	.	.	.	.	.	−3326400	.	.	.	.	
$(3)^3\,(2)\,(1)$	.	.	.	.	.	17740800	8870400	.	.	.	17740800	
$(3)^3\,(2)\,(1)$	.	.	.	.	.	.	−8870400	.	.	.	.	
$(3)^4$	.	.	.	.	.	.	.	.	.	.	.	
$(4)\,(1)^8$	−14968800	−7484400	−19958400	−9979200	−4989600	−3326400	−1663200	−9979200	−4989600	−2494800	−1663200	
$(4)\,(2)\,(1)^6$	44906400	29937600	59875200	39916800	24948000	19958400	11642400	59875200	34927200	19958400	14968800	
$(4)\,(2)^2\,(1)^4$	−44906400	−44906400	.	−29937600	−34927200	−29937600	−24948000	−29937600	−44906400	−39916800	−34927200	
$(4)\,(2)^3\,(1)^2$	14968800	29937600	.	.	14968800	.	14968800	.	14968800	29937600	14968800	
$(4)\,(2)^4$	32	.	.	.	.	.	.	.	.	.	.	
	128	−7484400	.	.	.	.	.	.	.	−7484400	.	
$(4)\,(3)\,(1)^5$	128	−64	−39916800	−19958400	−9979200	−13305600	−6652800	−79833600	−39916800	−19958400	−16632000	
$(4)\,(3)\,(2)\,(1)^3$	.	.	−12	19958400	19958400	39916800	26611200	.	39916800	39916800	59875200	
$(4)\,(3)\,(2)^2\,(1)$	.	.	−12	24	−9979200	.	−19958400	.	.	−19958400	−9979200	
$(4)\,(3)^2\,(1)^3$	48	.	−12	48	−48	−13305600	−6652800	.	.	.	−26611200	
	.	.	−24	72	.	−36	6652800	.	.	.	.	
$(4)\,(3)^2\,(2)$	72	.	−24	120	−144	−36	72	.	.	.	.	
$(4)^2\,(1)^4$	.	.	−32	.	.	.	.	29937600	14968800	7484400	4989600	
$(4)^2\,(2)\,(1)^2$	32	.	−32	64	.	.	.	16	−14968800	−14968800	−14968800	
$(4)^2\,(2)^2$	192	−64	−32	128	−128	.	.	16	−32	7484400	.	
$(4)^3\,(3)\,(1)$	48	.	−56	192	−48	−96	.	16	−64	64	9979200	
$(4)^3\,(2)$	192	−48	−96	384	−192	−192	.	16	−48	.	48	
$(5)\,(1)^7$	.	.	.	.	.	.	.	48	−192	96	192	
$(5)\,(2)\,(1)^5$	.	.	.	.	.	.	.	.	.	.	.	
$(5)\,(2)^3\,(1)^2$	.	.	.	.	.	.	.	.	.	.	.	
$(5)\,(2)^4\,(1)$	40	.	.	.	.	.	.	.	.	.	.	
$(5)\,(3)\,(1)^4$	.	.	−15	.	.	.	.	.	.	.	.	
$(5)\,(3)\,(2)\,(1)^2$	.	.	−15	30	.	.	.	.	.	.	.	
$(5)\,(3)\,(2)^2$	60	.	−15	60	−60	.	.	.	.	.	.	
$(5)\,(3)^2\,(1)$	.	.	−30	90	.	−45	.	.	.	.	.	
$(5)\,(4)\,(1)^3$	.	.	−40	20	.	.	.	20	.	.	.	
$(5)\,(4)\,(2)\,(1)$	60	.	−40	100	−40	.	.	20	−40	.	.	
$(5)\,(4)\,(3)$	90	.	−57	260	−150	−120	60	20	−60	.	60	
$(5)^2\,(1)^3$	.	.	−50	50	.	.	.	25	.	.	.	
$(5)^2\,(2)$	100	.	−50	150	−100	.	.	25	−50	.	.	
$(6)\,(1)^6$	.	.	.	.	.	.	.	.	.	.	.	
$(6)\,(2)\,(1)^4$	.	.	.	.	.	.	.	.	.	.	.	
$(6)\,(2)^2\,(1)^2$	24	.	.	.	.	.	.	.	.	.	.	
$(6)\,(2)^3$	120	−48	.	.	.	.	.	.	.	.	.	
$(6)\,(3)\,(1)^3$	.	.	−18	18	.	.	.	.	.	.	.	
$(6)\,(3)\,(2)\,(1)$	36	.	−18	54	−36	.	.	.	.	.	.	
$(6)\,(3)^2$	54	.	−36	144	−108	−54	54	.	.	.	.	
$(6)\,(4)\,(1)^2$	20	.	−48	72	.	−12	.	24	−24	.	.	
$(6)\,(4)\,(2)$	104	−40	−48	168	−144	−12	24	24	−72	48	.	
$(6)\,(5)\,(1)$	40	.	−60	120	−30	−15	.	30	−30	.	.	
$(6)^2$	132	−24	−72	216	−144	−36	36	36	−72	36	.	
$(7)\,(1)^5$	.	.	−7	.	.	.	.	.	.	.	.	
$(7)\,(2)\,(1)^3$	.	.	−7	14	.	.	.	.	.	.	.	
$(7)\,(2)^2\,(1)$	36	.	−7	28	−28	.	.	.	.	.	.	
$(7)\,(3)\,(1)^2$	.	.	−28	63	.	−21	.	.	.	.	.	
$(7)\,(3)\,(2)$	84	.	−28	119	−126	−21	42	.	.	.	.	
$(7)\,(4)\,(1)$	56	.	−63	168	−42	−36	.	28	−56	.	28	
$(7)\,(5)$	105	.	−77	245	−140	−70	35	35	−70	.	35	
$(8)\,(1)^4$	.	.	−16	.	.	.	.	4	.	.	.	
$(8)\,(2)\,(1)^2$	16	.	−16	32	.	.	.	4	−8	.	.	
$(8)\,(2)^2$	128	−32	−16	64	−64	.	.	4	−16	16	.	
$(8)\,(3)\,(1)$	24	.	−40	120	−24	−48	.	4	−12	.	12	
$(8)\,(4)$	144	−24	−80	288	−160	−112	32	36	−112	40	80	
$(9)\,(1)^3$	.	.	−27	18	.	.	.	9	.	.	.	
$(9)\,(2)\,(1)$	54	.	−27	72	−36	.	.	9	−18	.	.	
$(9)\,(3)$	81	.	−54	207	−135	−81	54	9	−27	.	27	
$(10)\,(1)^2$	10	.	−40	60	.	−10	.	15	−10	.	.	
$(10)\,(2)$	130	−20	−40	140	−120	−10	20	15	−40	20	.	
$(11)\,(1)$	44	.	−55	132	−33	−33	.	22	−33	.	11	
(12)	120	−12	−72	240	−144	−72	36	30	−72	18	36	

Table 5.12 (*cont.*)

$w=12$ (iv)	a_4^3	$a_4a_7^2$	$a_4a_6a_1^2$	$a_4a_5^2a_2^2$	$a_4a_5^2a_1$	$a_4a_5a_3$	$a_4a_4^2a_2$	$a_4a_4a_1^3$	$a_4a_4a_3a_1$	$a_4a_4a_2^2$
(1)^{12}	34650	3991680	1995840	997920	498960	665280	332640	166320	110880	166320	83160
(2) (1)^{10}	−623700	−3991680	−21954240	−11975040	−6486480	−8648640	−4656960	−2494800	−1774080	−2661120	−1413720
(2)^2 (1)^8	4054050	59875200	49896000	35925120	23950080	29937600	19293120	11975040	9313920	12972960	7817040
(2)^3 (1)^6	−11226600	.	−29937600	−39916800	−37920960	−29937600	−20937600	−14615360	−10958400	−10958400	−10465680
(2)^4 (1)^4	12162150	.	.	14968800	27442800	.	14968800	22453200	14968800	7484400	13721400
(2)^5 (1)^2	−5613300	.	.	.	−7484400	.	.	−7484400	.	.	−3742200
(2)^6	935550	.	.	.	.	.	.	.	.	.	.
(3) (1)^9	831600	79833600	39916800	19958400	9979200	14636160	7318080	3659040	2661120	4656960	2328480
(3) (2) (1)^7	−9979200	−79833600	−79833600	−59875200	−39916800	−66528000	−40582080	−23950080	−21288960	−36590400	−20023680
(3) (2)^2 (1)^5	34927200	.	39916800	59875200	59875200	59875200	63201600	51891840	53222400	49896000	43243200
(3) (2)^3 (1)^3	−20937600	.	.	.	.	.	.	.	.	.	.
(3) (2)^4 (1)	7484400	.	.	−19958400	−39916800	.	−29937600	−46569600	−39916800	−9979200	−29937600
(3)^2 (1)^6	6652800	.	.	.	9979200	.	.	14968800	.	.	4989600
(3)^2 (2) (1)^4	−39916800	.	.	.	.	26611200	13305600	6652800	9313920	26611200	13305600
(3)^2 (2)^2 (1)^2	19958400	.	.	.	.	−26611200	−26611200	−19958400	−39916800	−26611200	−26611200
(3)^2 (2)^3	.	.	.	.	.	.	13305600	19958400	33264000	.	13305600
(3)^3 (1)^3	17740800	.	.	.	.	.	.	−6652800	.	.	.
(3)^3 (2) (1)	.	.	.	.	.	.	.	.	8870400	.	.
(3)^4	.	.	.	.	.	.	.	.	−8870400	.	.
(4) (1)^8	−623700	−11975040	−59875200	−29937600	−14968800	−19958400	−9979200	−4989600	−3326400	−5987520	−2993760
(4) (2) (1)^6	7484400	.	59875200	59875200	44906400	59875200	39916800	24948000	19958400	39916800	22952160
(4) (2)^2 (1)^4	−26195400	.	.	−29937600	−44906400	.	−29937600	−34927200	−29937600	−29937600	−34927200
(4) (2)^3 (1)^2	22453200	.	.	.	14968800	.	.	14968800	.	.	14968800
(4) (2)^4	−5613300	.	.	.	.	.	.	.	.	.	.
(4) (3) (1)^5	−9979200	.	.	.	.	−39916800	−19958400	−9979200	−13305600	−59875200	−29937600
(4) (3) (2) (1)^3	59875200	.	.	.	.	.	19958400	19958400	39916800	19958400	39916800
(4) (3)^2 (1)	−29937600	.	.	.	.	.	.	−9979200	.	.	−9979200
(4) (3)^2 (2)	−39916800	.	.	.	.	.	.	.	−13305600	.	.
(4) (3)^3	.	.	.	.	.	.	.	.	.	29937600	14968800
(4)^2 (1)^4	3742200	.	.	.	.	.	.	.	.	.	.
(4)^2 (2) (1)^2	−22453200	.	.	.	.	.	.	.	.	.	−14968800
(4)^2 (2)^2	11226600	.	.	.	.	.	.	.	.	.	.
(4)^2 (3) (1)	29937600	.	.	.	.	.	.	.	.	.	.
(4)^2	−7484400	.	.	.	.	.	.	.	.	.	.
	−64	.	.	.	.	.	.	.	.	.	.
(5) (1)^7	.	95800320	47900160	23950080	11975040	15966720	7983360	3991680	2661120	3991680	1995840
	.	5	.	.	.	.	.	.	.	.	.
(5) (2) (1)^5	.	.	−47900160	−47900160	−35925120	−47900160	−31933440	−19958400	−15966720	−23950080	−13970880
	.	5	−10	.	.	.	.	.	.	.	.
(5) (2)^2 (1)^3	.	.	.	23950080	35925120	.	23950080	27941760	23950080	11975040	17962560
	.	5	−20	20	.	.	.	.	.	.	.
(5) (2)^3 (1)	.	.	.	.	−11975040	.	.	−11975040	.	.	−5987520
	.	5	−30	60	−40	.	.	.	.	.	.
(5) (3) (1)^4	.	.	.	.	.	31933440	15966720	7983360	10644480	31933440	15966720
	.	5	−15	.	.	15	.	.	.	.	.
(5) (3) (2)^2	.	.	.	.	.	−15966720	−15966720	−31933440	.	.	−15966720
	.	5	−25	30	.	15	−30	.	.	.	.
(5) (3)^2 (1)	.	.	.	.	.	.	7983360	.	.	.	.
	.	5	−35	80	−60	15	−60	60	.	.	.
(5) (4) (1)^3	.	.	.	.	.	.	.	.	10644480	.	.
	.	5	−30	45	.	30	−90	.	45	.	.
(5) (4) (2) (1)	.	.	.	.	.	.	.	.	.	−23950080	−11975040
	.	5	−20	10	.	20	.	.	.	−20	.
(5) (4) (3)	.	.	.	.	.	.	.	.	.	.	11975040
	.	5	−30	50	−20	20	−40	.	.	.	.
(5)^2 (1)^2	.	5	−35	70	−30	35	−120	30	.	−20	40
	.	10	−50	50	.	50	−50	.	60	−20	60
(5)^2 (2)	.	10	−70	150	−100	50	−150	100	.	−50	.
(6) (1)^6	.	6	.	.	.	.	.	.	.	.	100
(6) (2) (1)^4	.	6	−12	.	.	.	.	.	.	.	.
(6) (2)^2 (1)^2	.	6	−24	24	.	.	.	.	.	.	.
(6) (2)^3	.	.	.	.	.	.	.	.	.	.	.
(6) (3) (1)^3	.	6	−36	72	−48	.	.	.	.	.	.
(6) (3) (2) (1)	.	6	−18	.	.	.	.	.	.	.	.
(6) (3)^2	.	6	−30	36	.	18	.	.	.	.	.
(6) (4) (1)^2	.	6	−36	54	.	18	−36	.	.	.	.
(6) (4) (2)	.	6	−24	12	.	36	−108	.	54	.	.
(6) (5) (1)	.	6	−36	60	−24	24	.	.	.	−24	.
(6)^2	.	11	−60	75	−10	24	−48	.	.	−24	48
(7) (1)^5	.	12	−72	108	−24	60	−90	.	15	−60	30
(7) (2) (1)^3	.	7	−7	.	.	72	−144	.	30	−72	72
	.	7	−21	14	.	.	.	.	.	.	.
(7) (2)^2 (1)	.	.	.	.	.	.	.	.	.	.	.
(7) (3) (1)^2	.	7	−35	56	−28	.	.	.	.	.	.
(7) (3) (2)	.	7	−28	21	.	.	.	.	.	.	.
(7) (4) (1)	.	7	−42	77	−42	21	−21	.	.	.	.
(7) (5)	.	7	−35	42	−14	21	−63	42	.	.	.
(8) (1)^4	.	12	−77	140	−30	28	−28	.	.	.	.
(8) (2) (1)^2	.	8	−16	.	.	70	−175	70	.	−28	28
(8) (2)^2	.	8	−32	32	.	8	.	.	35	−70	105
(8) (3) (1)	.	8	−48	96	−64	8	−16	.	.	.	.
(8) (4)	−16	8	−40	48	.	8	−32	32	.	.	.
	.	8	−48	80	−32	32	−72	.	.	.	.
(9) (1)^3	.	.	.	.	.	40	−96	16	24	.	.
(9) (2) (1)	.	9	−27	9	.	18	.	.	32	−32	64
(9) (3)	.	9	−45	63	−18	18	−36	.	.	−9	.
(10) (1)^2	.	9	−54	90	−27	45	−135	27	.	−9	18
(10) (2)	.	10	−40	30	.	30	−120	.	54	−9	27
(11) (1)	.	10	−60	110	−60	30	−80	40	.	−20	.
(12)	−4	11	−55	66	−11	44	−66	.	.	−20	40
	.	12	−72	120	−48	60	−144	36	11	−33	22
	.	.	.	.	.	.	.	.	36	−48	72

Table 5.12 (*cont.*)

$w=12$ (v)	$a_1a_4a_5$	$a_1^2a_3^2$	$a_2^3a_3$	$a_4a_5^2$	$a_1a_2a_3^2$	$a_1a_3^2a_5^2$	$a_4a_5^3$	$a_4a_6a_7^2$	$a_1a_3a_5a_7$	$a_4a_5^2$	$a_3a_4a_7^2$
$(1)^{12}$	27720	33264	16632	665280	332640	166320	83160	110880	55440	18480	27720
$(2)\ (1)^{10}$	−526680	−665280	−349272	−9979200	−5322240	−2827440	−1496880	−1995840	−1053360	−388080	−582120
$(2)^2\ (1)^8$	3492720	4324320	2494800	29937600	19958400	12640320	7733880	9979200	5987520	2661120	3825360
$(2)^3\ (1)^6$	−9812880	−9979200	−7151760	−9979200	−19958400	−14209680	−10209360	−13632000	−13305600	−7761600	−9147600
$(2)^4\ (1)^4$	11226600	7484400	8731800	.	4989600	12474000	16216200	4989600	10810800	9147600	6237000
$(2)^5\ (1)^3$	−3742200	.	−3742200	.	.	−2494800	−7484400	.	−2494800	−2494800	−1247400
$(2)^6$	.	.	.	.	.	.	1247400	.	.	.	.
$(3)\ (1)^9$	831600	1330560	665280	26611200	13305600	6652800	3326400	4656960	2328480	831120	1330560
$(3)\ (2)\ (1)^7$	−9313920	−14636160	−7983360	−79833600	−53222400	−33264000	−19958400	−29937600	−17297280	−7983360	−13305600
$(3)\ (2)^2\ (1)^5$	30935520	33264000	23950080	.	39916800	46569600	39916800	49896000	39916800	26611200	33264000
$(3)\ (2)^3\ (1)^3$	−33264000	−19958400	−26611200	.	.	−19958400	−33264000	−3326400	−26611200	−31046400	−13305600
$(3)\ (2)^4\ (1)$	7484400	.	9979200	.	.	.	9979200	.	1663200	3326400	.
$(3)^2\ (1)^6$	5987520	13305600	6652800	26611200	13305600	6652800	3326400	13305600	6652800	3769920	9979200
$(3)^2\ (2)\ (1)^4$	−29937600	−26611200	−19958400	.	−13305600	−13305600	−9979200	−3991680	−26611200	−23284800	−33264000
$(3)^2\ (2)^2\ (1)^2$	29937600	13305600	19958400	.	.	6652800	9979200	.	19958400	36590400	33264000
$(3)^2\ (2)^3$	−3326400	.	−6652800	.	.	.	−3326400	.	.	−1108800	.
$(3)^3\ (1)^3$	8870400	.	.	.	.	.	.	8870400	4435200	5913600	8870400
$(3)^3\ (2)\ (1)$	−8870400	.	.	.	.	.	.	.	−4435200	−17740800	.
$(3)^4$	.	.	.	.	.	.	.	.	.	2956800	.
$(4)\ (1)^8$	−997920	−1995840	−997920	−59875200	−29937600	−14968800	−7484400	−9979200	−9989600	−1663200	−2661120
$(4)\ (2)\ (1)^6$	9646560	19958400	10977120	59875200	.	−29937600	−44906400	−44906400	−29937600	−34927200	−24948000
$(4)\ (2)^2\ (1)^4$	−24948000	−29937600	−24948000	.	−29937600	.	14968800	29937600	.	14968800	14968800
$(4)\ (2)^3\ (1)^2$	14968800	.	14968800	.	.	14968800	.	.	−19958400	−9979200	−6652800
$(4)\ (2)^4$	.	.	.	.	.	.	−7484400	.	19958400	19958400	26611200
$(4)\ (3)\ (1)^5$	−11975040	−39916800	−19958400	.	.	.	.	.	.	−9979200	−19958400
$(4)\ (3)\ (2)\ (1)^3$	46569600	39916800	39916800	.	.	.	.	.	.	.	.
$(4)\ (3)\ (2)^2\ (1)$	−19958400	.	−19958400	.	.	.	.	.	.	.	.
$(4)\ (3)^2\ (1)^2$	−19958400	.	.	.	.	.	.	.	.	.	.
$(4)\ (3)^2\ (2)$	6652800	.	.	.	.	.	.	.	.	.	.
$(4)^2\ (1)^4$	4989600	27937600	14968800	.	.	.	.	.	.	.	.
$(4)^2\ (2)\ (1)^2$	−14968800	.	−14968800	.	.	.	.	.	.	.	.
$(4)^2\ (2)^2$	.	.	.	.	.	.	.	.	.	.	.
$(4)^3\ (1)^2$	9979200	.	.	.	.	.	.	.	.	.	.
$(4)^3$	.	.	.	.	.	.	.	.	.	.	.
$(5)\ (1)^7$	665280	1596672	798336	95800320	47900160	23950080	11975040	15966720	7983360	2661120	3991680
$(5)\ (2)\ (1)^5$	−5987520	−15966720	−8781696	.	−47900160	−47900160	−35925120	−47900160	−31933440	−15966720	−23950080
$(5)\ (2)^2\ (1)^3$	13970880	23950080	1958400	.	.	23950080	35925120	.	23950080	23950080	11975040
$(5)\ (2)^3\ (1)$	−5987520	.	−11975040	.	.	.	11975040	.	23950080	.	.
$(5)\ (3)\ (1)^4$	6652800	31933440	15966720	.	.	.	.	31933440	15966720	10644480	31933440
$(5)\ (3)\ (2)\ (1)^2$	−23950080	−31933440	−31933440	.	.	.	.	.	−15966720	−31933440	.
$(5)\ (3)\ (2)^2$	3991680	.	15966720	.	.	.	.	.	.	.	.
$(5)\ (3)^2\ (1)$	10644480	.	.	.	.	.	.	.	.	10644480	.
$(5)\ (4)\ (1)^3$	−3991680	−47900160	−23950080	.	.	.	.	.	.	.	−23950080
$(5)\ (4)\ (2)\ (1)$	11975040	.	23950080	.	.	.	.	.	.	.	.
$(5)\ (4)\ (3)$	−7983360	.	.	.	.	.	.	.	.	.	.
	−60										
$(5)^2\ (1)^3$	.	19160064	9580032	.	.	.	.	.	.	.	.
	.	25									
$(5)^2\ (2)$	.	.	−9580032	.	.	.	.	.	.	.	.
	.	25	−50								
$(6)\ (1)^6$	.	.	.	−79833600	−39916800	−19958400	−9979200	−13305600	−6652800	−2217600	−3326400
	.	.	.	−6							
$(6)\ (2)\ (1)^4$	.	.	.	.	39916800	39916800	29937600	39916800	26611200	13305600	19958400
	.	.	.	−6	12						
$(6)\ (2)^2\ (1)^2$	.	.	.	.	.	−19958400	−29937600	.	−19958400	−19958400	−9979200
	.	.	.	−6	24	−24					
$(6)\ (2)^3$	.	.	.	.	.	.	9979200	.	.	.	.
	.	.	.	−6	36	−72	48				
$(6)\ (3)\ (1)^3$	.	.	.	.	.	.	.	−26611200	−13305600	−8870400	−26611200
	.	.	.	−6	18	.	.	−18			
$(6)\ (3)\ (2)\ (1)$	.	.	.	.	.	.	.	.	13305600	26611200	.
	.	.	.	−6	30	−36	.	−18	36		
$(6)\ (3)^2$	.	.	.	.	.	.	.	.	.	−8870400	.
	.	.	.	−6	36	−54	.	−36	108	−54	
$(6)\ (4)\ (1)^2$	.	.	.	.	.	.	.	.	.	.	19958400
$(6)\ (4)\ (2)$	.	.	.	−6	24	−12	.	−24	.	.	24
$(6)\ (5)\ (1)$	.	.	.	−6	36	−60	24	−24	48	.	24
$(6)^2$	.	30	.	−6	30	−30	.	−30	30	.	30
$(7)\ (1)^5$	.	36	.	−12	72	−108	24	−72	144	−36	72
$(7)\ (2)\ (1)^3$	.	.	.	−7	.	.	.	.	.	.	.
	.	.	.	−7	14	.	.	.	.	.	.
$(7)\ (2)^2\ (1)$	.	.	.	.	.	.	.	.	.	.	.
$(7)\ (3)\ (1)^2$	.	.	.	−7	28	−28	.	.	.	.	.
$(7)\ (3)\ (2)$	.	.	.	−7	21	.	.	−21	.	.	.
$(7)\ (4)\ (1)$	.	.	.	−7	35	−42	.	−21	42	.	.
$(7)\ (5)$	−35	35	−35	−7	28	−14	.	−28	.	.	28
$(8)\ (1)^4$	.	.	.	−7	35	−35	.	−35	35	.	35
$(8)\ (2)\ (1)^2$	.	.	.	−8	8	.	.	.	.	.	.
$(8)\ (2)^2$	.	.	.	−8	24	−16	.	.	.	.	.
$(8)\ (3)\ (1)$	.	.	.	−8	40	−64	32	.	.	.	.
$(8)\ (4)$	−32	.	.	−8	32	−24	.	−24	24	.	.
	.	.	.	−8	40	−48	16	−32	32	.	32
$(9)\ (1)^3$	.	.	.	−9	18	.	.	−9	.	.	.
$(9)\ (2)\ (1)$	.	.	.	−9	36	−36	.	−9	18	.	.
$(9)\ (3)$	−27	.	.	−9	45	−54	.	−36	81	−27	.
$(10)\ (1)^2$	.	5	.	−10	30	−10	.	−20	.	.	10
$(10)\ (2)$	.	5	−10	−10	50	−70	20	−20	40	.	10
$(11)\ (1)$	.	11	.	−11	44	−33	.	−33	22	.	22
(12)	−24	18	−12	−12	60	−72	12	−48	72	−12	36

Table 5.12 (*cont.*)

$w-12$ (vi)	$a_1a_4a_7$	$a_1a_5a_7$	a_4^2	$a_7a_1^2$	$a_7a_2a_1^2$	$a_7a_1^2a_1$	$a_7a_2a_1^2$	$a_7a_2a_3$	$a_7a_4a_1$	a_7a_5	$a_8a_1^2$
$(1)^{12}$	13860	5544	924	95040	47520	23760	15840	7920	3960	792	11880
$(2)\,(1)^{10}$	−304920	−138600	−27720	−1995840	−1045440	−546480	−380160	−198000	−106920	−24552	−332640
$(2)^2\,(1)^8$	2203740	1164240	291060	9970200	5987520	3516480	2661120	1520640	926640	261360	2494800
$(2)^3\,(1)^6$	−6486480	−3825360	−1275120	−9970200	−9970200	−7983360	−6652800	−4656960	−3160080	−1164240	−4989600
$(2)^4\,(1)^4$	7692300	4573800	2286900	.	4989600	7484400	4989600	5821200	3742200	2079000	1247400
$(2)^5\,(1)^2$	−3742200	−1247400	−1247400	.	.	−2494800	.	−2494800	−1247400	−1247400	.
$(2)^6$	623700	.	207900	.	.	.	.	.	.	.	.
$(3)\,(1)^9$	665280	332640	73920	6652800	3326400	1663200	1140480	570240	308880	71280	1330560
$(3)\,(2)\,(1)^7$	−7318080	−4656960	−1330560	−39916800	−23284800	−13305600	−10644480	−5802480	−3991680	−1235520	−13305600
$(3)\,(2)^2\,(1)^5$	23284800	16632000	6652800	199584000	29937600	26611200	26611200	18627840	14968800	6320160	19958400
$(3)\,(2)^3\,(1)^3$	−23284800	−16632000	−11088000	.	−9970200	−19958400	−13305600	−19958400	−13305600	−9970200	.
$(3)\,(2)^4\,(1)$	6652800	1663200	3326400	.	.	4989600	.	6652800	2494800	4158000	.
$(3)^2\,(1)^6$	4689600	4656960	1552320	26611200	13305600	6652800	6652800	3326400	3326400	1330560	13305600
$(3)^2\,(2)\,(1)^4$	−21621600	−19958400	−9979200	.	−13305600	−13305600	−26611200	−16632000	−19958400	−9979200	−13305600
$(3)^2\,(2)^2\,(1)^2$	18295200	16632000	16632000	.	.	6652800	6652800	16632000	9979200	13305600	.
$(3)^2\,(2)^3$	−1663200	.	−11088800	.	.	.	.	−3326400	.	−3326400	.
$(3)^3\,(1)^3$	4435200	4435200	2956800	.	.	.	8870400	4435200	8870400	4435200	.
$(3)^3\,(2)\,(1)$	−4435200	−4435200	−8870400	.	.	.	.	−4435200	.	−4435200	.
$(3)^4$	.	.	1478400	.	.	.	.	.	.	.	.
$(4)\,(1)^8$	−1330560	−665280	−166320	−19958400	−9979200	−4989600	−3326400	−1663200	−855360	−190080	−4989600
$(4)\,(2)\,(1)^6$	11309760	7983360	2661120	59875200	39916800	24948000	19958400	11642400	7983360	2661120	29937600
$(4)\,(2)^2\,(1)^4$	−24948000	−19958400	−9979200	.	−29937600	−34927200	−29937600	−24948000	−19958400	−9979200	−14968800
$(4)\,(2)^3\,(1)^2$	19958400	9979200	9979200	.	.	14968800	.	14968800	9979200	9979200	.
$(4)\,(2)^4$	−4989600	.	−2494800	.	.	.	.	.	.	.	.
$(4)\,(3)\,(1)^5$	−13305600	−16632000	−6652800	−39916800	−19958400	−9979200	−13305600	−6652800	−9979200	−5322240	−39916800
$(4)\,(3)\,(2)\,(1)^3$	33264000	39916800	26611200	.	19958400	19958400	39916800	26611200	39916800	26611200	.
$(4)\,(3)\,(2)^2\,(1)$	−19958400	−9979200	−19958400	.	.	−9979200	.	−19958400	−9979200	−19958400	.
$(4)\,(3)^2\,(1)^2$	−3326400	−6652800	−6652800	.	.	.	−13305600	−6652800	−19958400	−13305600	.
$(4)\,(3)^2\,(2)$	3326400	.	6652800	.	.	.	.	6652800	.	6652800	.
$(4)^2\,(1)^4$	7484400	14968800	7484400	.	.	.	.	.	4989600	4989600	14968800
$(4)^2\,(2)\,(1)^2$	−14968800	−14968800	−14968800	.	.	.	.	.	−14968800	−14968800	.
$(4)^2\,(2)^2$	7484400	.	7484400	.	.	.	.	.	.	.	.
$(4)^2\,(3)\,(1)$	.	.	.	.	.	.	.	.	9979200	9979200	.
$(4)^3$	.	.	.	.	.	.	.	.	.	.	.
$(5)\,(1)^7$	1995840	931392	266112	47900160	23950080	11975040	7983360	3991680	1995840	418176	15966720
$(5)\,(2)\,(1)^5$	−13970880	−9979200	−3991680	−47900160	−47900160	−35925120	−31933440	−19958400	−13970880	−4790016	−47900160
$(5)\,(2)^2\,(1)^3$	17962560	17962560	11975040	.	23950080	35925120	23950080	27941760	17962560	11975040	.
$(5)\,(2)^3\,(1)$	−5987520	−1995840	−3991680	.	.	−11975040	.	−11975040	−5987520	−7983360	.
$(5)\,(3)\,(1)^4$	15966720	21288960	10644480	.	.	.	15966720	7983360	15966720	9313920	31933440
$(5)\,(3)\,(2)\,(1)^2$	−15966720	−31933440	−31933440	.	.	.	−15966720	−15966720	−15966720	−23950080	.
$(5)\,(3)\,(2)^2$	.	.	.	.	.	.	.	7983360	.	11975040	.
$(5)\,(3)^2\,(1)$	.	5322240	10644480	.	.	.	.	.	.	5322240	.
$(5)\,(4)\,(1)^3$	−11975040	−35925120	−23950080	.	.	.	.	.	−11975040	−15966720	.
$(5)\,(4)\,(2)\,(1)$	11975040	11975040	23950080	.	.	.	.	.	11975040	23950080	.
$(5)\,(4)\,(3)$	.	.	.	.	.	.	.	.	.	−7983360	.
$(5)^2\,(1)^2$	.	19160064	19160064	.	.	.	.	.	.	9580032	.
$(5)^2\,(2)$	.	.	.	.	.	.	.	.	.	−9580032	.
$(6)\,(1)^6$	−1663200	−665280	−221760	−79833600	−39916800	−19958400	−13305600	−6652800	−3326400	−665280	−39916800
$(6)\,(2)\,(1)^4$	11642400	6652800	3326400	.	39916800	39916800	39916800	26611200	19958400	6652800	39916800
$(6)\,(2)^2\,(1)^2$	−14968800	−9979200	−9979200	.	.	−19958400	.	−19958400	−9979200	−9979200	.
$(6)\,(2)^3$	4989600	.	3326400	.	.	.	.	.	.	.	.
$(6)\,(3)\,(1)^3$	−13305600	−13305600	−8870400	.	.	.	−26611200	−13305600	−26611200	−13305600	.
$(6)\,(3)\,(2)\,(1)$	13305600	13305600	26611200	.	.	.	.	13305600	.	13305600	.
$(6)\,(3)^2$	.	.	−8870400	.	.	.	.	.	.	.	.
$(6)\,(4)\,(1)^2$	9979200	19958400	19958400	.	.	.	.	.	19958400	19958400	.
$(6)\,(4)\,(2)$	−9979200	.	−19958400	.	.	.	.	.	.	.	.
$(6)\,(5)\,(1)$	−48	.	.	.	.	.	.	.	.	.	.
	.	−15966720	−31933440	.	.	.	.	.	.	−15966720	.
$(6)^2$	.	−30	13305600	.	.	.	.	.	.	.	.
$(7)\,(1)^5$	−72	−72	36	68428800	34214400	17107200	11404800	5702400	2851200	570240	68428800
$(7)\,(2)\,(1)^3$	.	.	.	7	−34214400	−34214400	−34214400	−22809600	−17107200	−5702400	.
$(7)\,(2)^2\,(1)$	.	.	.	7	−14	17107200	.	17107200	8553600	8553600	.
$(7)\,(3)\,(1)^3$	.	.	.	7	−28	28	22809600	11404800	22809600	11404800	.
$(7)\,(3)\,(2)$	.	.	.	7	−21	.	21	−11404800	.	−11404800	.
$(7)\,(4)\,(1)$	.	.	.	7	−35	42	21	−42	−17107200	−17107200	.
$(7)\,(5)$	.	−35	.	7	−35	35	35	−35	−35	13685760	.
$(8)\,(1)^4$	.	.	.	8	.	.	.	.	.	35	−59873200
$(8)\,(2)\,(1)^2$	.	.	.	8	−16	.	.	.	.	.	−8
$(8)\,(2)^2$	.	.	.	8	−32	32	.	.	.	.	−8
$(8)\,(3)\,(1)$	.	.	.	8	−24	.	24	.	.	.	−8
$(8)\,(4)$	−32	.	.	8	−32	16	32	.	−32	.	−8
$(9)\,(1)^3$	.	.	.	9	−9	.	.	.	.	.	−9
$(9)\,(2)\,(1)$	.	.	.	9	−27	18	.	.	.	.	−9
$(9)\,(3)$	.	.	.	9	−36	27	27	−27	.	.	−9
$(10)\,(1)^2$	.	.	.	10	−20	.	10	.	.	.	−10
$(10)\,(2)$	−20	.	.	10	−40	40	10	−20	.	.	−10
$(11)\,(1)$	.	−11	.	11	−33	11	22	.	−11	.	−11
(12)	−24	−24	6	12	−48	36	36	−24	−24	12	−12

Table 5.12 (*cont.*)

$w=12$ (vii)	$a_1a_2a_3^3$	$a_2a_3^3$	$a_1a_2a_3$	a_2a_4	$a_3a_3^3$	$a_2a_3a_3$	a_2a_5	$a_{10}a_3^3$	$a_{10}a_3$	$a_{11}a_1$	a_{12}
$(1)^{12}$	5940	2970	1980	495	1320	660	220	132	66	12	1
$(2)\,(1)^{10}$	−172260	−89100	−61380	−16830	−47520	−24420	−8580	−5940	−3036	−660	−66
$(2)^2\,(1)^8$	1413720	702990	582120	188595	498960	273240	106920	83160	44550	11880	1485
$(2)^3\,(1)^6$	−3742200	−2577960	−2079000	−873180	−1663200	−1081080	−526680	−415800	−249480	−83160	−13860
$(2)^4\,(1)^4$	3118500	3430350	2702700	1611225	1247400	1455300	1039500	623700	519750	207900	51975
$(2)^5\,(1)^2$	−623700	−1871100	−623700	−935550	.	−623700	−623700	−124740	−374220	−124740	−62370
$(2)^6$	.	311850	.	155925	.	.	.	.	62370	.	10395
$(3)\,(1)^9$	665280	332640	225720	59400	221760	110880	37400	31680	15840	3960	440
$(3)\,(2)\,(1)^7$	−7318680	−3991680	−2093760	−997920	−3326400	−1774080	−681120	−665280	−348480	−110880	−15840
$(3)\,(2)^2\,(1)^5$	16632000	11975040	10810800	5155920	9979200	6652800	3492720	3326400	1995840	831600	166320
$(3)\,(2)^3\,(1)^3$	−9979200	−13305600	−11642400	−8316000	−3326400	−6652800	−6652800	−3326400	−3326400	−1663200	−554400
$(3)\,(2)^4\,(1)$	.	4989600	415800	2910600	.	1663200	2079000	.	1663200	415800	415800
$(3)^2\,(1)^6$	6652800	3326400	2661120	997920	4435200	2217600	813120	1108800	554400	221760	36960
$(3)^2\,(2)\,(1)^4$	−13305600	−9979200	−13305600	−8316000	−13305600	−8870400	−5544000	−6652800	−3880800	−2217600	−554400
$(3)^2\,(2)^2\,(1)^2$	6652800	9979200	13305600	11642400	.	6652800	9979200	3326400	4989600	3326400	1663200
$(3)^2\,(2)^3$	.	−3326400	.	−1663200	.	.	−1108800	.	−1663200	.	−554400
$(3)^3\,(1)^3$	.	.	4435200	4435200	2956800	1478400	1971200	2956800	1478400	1478400	492800
$(3)^3\,(2)\,(1)$	.	.	−4435200	−4435200	.	−1478400	−5913600	.	−1478400	−1478400	−1478400
$(3)^4$	.	.	.	.	.	.	985600	.	.	.	246400
$(4)\,(1)^8$	−2494800	−1247400	−831600	−210870	−997920	−498960	−166320	−166320	−83160	−23760	−2970
$(4)\,(2)\,(1)^6$	17463600	9979200	7484400	2577960	9979200	5488560	2162160	2494800	1330560	498960	83160
$(4)\,(2)^2\,(1)^4$	−22453200	−19958400	−17463600	−9355500	−14968800	−12474000	−7484400	−7484400	−4989600	−2494800	−623700
$(4)\,(2)^3\,(1)^2$	7484400	14968800	7484400	8731800	.	7484400	7484400	2494800	4989600	2494800	1247400
$(4)\,(2)^4$	.	−3742200	.	−2182950	.	.	.	.	−1247400	.	−311850
$(4)\,(3)\,(1)^5$	−19958400	−9979200	−8316000	−3659040	−19958400	−9979200	−3659040	−6652800	−3326400	−1663200	−332640
$(4)\,(3)\,(2)\,(1)^3$	19958400	1958400	29937600	23284800	19958400	19958400	16632000	19958400	13305600	9979200	3326400
$(4)\,(3)\,(2)^2\,(1)$	.	−9979200	−4989600	−14968800	.	−9979200	−14968800	.	−9979200	−4989600	−4989600
$(4)\,(3)^2\,(1)^2$	.	.	−13305600	−16632000	.	.	−6652800	−6652800	−3326400	−6652800	−3326400
$(4)\,(3)^2\,(2)$	.	.	.	3326400	.	.	6652800	.	3326400	.	3326400
$(4)^2\,(1)^4$	7484400	3742200	2494800	1871100	14968800	7484400	2494800	7484400	3742200	2494800	623700
$(4)^2\,(2)\,(1)^2$	−7484400	−7484400	−7484400	−11226600	.	−7484400	−7484400	−7484400	−7484400	−7484400	−3742200
$(4)^2\,(2)^2$	.	3742200	.	5613300	.	.	.	.	3742200	.	1871100
$(4)^2\,(3)\,(1)$	.	.	4989600	14968800	.	.	4989600	.	.	4989600	4989600
$(4)^3$	.	.	.	3742200	.	.	.	.	.	.	1247400
$(5)\,(1)^7$	7983360	3991680	2661120	665280	3991680	1995840	665280	798336	399168	133056	19008
$(5)\,(2)\,(1)^5$	−31933440	−19958400	−15966720	−5987520	−23950080	−13970880	−5987520	−7983360	−4390848	−1995840	−399168
$(5)\,(2)^2\,(1)^3$	23950080	27941760	23950080	13970880	11975040	17962560	13970880	11975040	9979200	5987520	1995840
$(5)\,(2)^3\,(1)$	.	−11975040	.	−5987520	.	−5987520	−5987520	.	−5987520	−1995840	−1995840
$(5)\,(3)\,(1)^4$	15966720	7983360	10644480	6652800	31933440	15966720	6652800	15966720	7983360	5322240	1330560
$(5)\,(3)\,(2)\,(1)^2$	−15966720	−15966720	−31933440	−23950080	.	−15966720	−23950080	−15966720	−15966720	−15966720	−7983360
$(5)\,(3)\,(2)^2$	.	7983360	.	3991680	.	.	3991680	.	7983360	.	3991680
$(5)\,(3)^2\,(1)$	.	.	10644480	10644480	.	.	10644480	.	.	5322240	5322240
$(5)\,(4)\,(1)^3$	.	.	.	−3991680	−23950080	−11975040	−3991680	−23950080	−11975040	−11975040	−3991680
$(5)\,(4)\,(2)\,(1)$	.	.	.	11975040	.	11975040	11975040	.	11975040	11975040	11975040
$(5)\,(4)\,(3)$	.	.	.	−7983360	.	.	−7983360	.	.	.	−7983360
$(5)^2\,(1)^2$	.	.	.	.	.	.	.	9580032	4790016	9580032	4790016
$(5)^2\,(2)$	.	.	.	.	.	.	.	.	−4790016	.	−4790016
$(6)\,(1)^6$	−19958400	−9979200	−6652800	−1663200	−13305600	−6652800	−2217600	−3326400	−1663200	−6652800	−110880
$(6)\,(2)\,(1)^4$	39916800	29937600	26611200	11642400	39916800	26611200	13305600	19958400	11642400	6652800	1663200
$(6)\,(2)^2\,(1)^2$	−19958400	−29937600	−19958400	−14968800	.	−19958400	−19958400	−9979200	−14968800	−9979200	−4989600
$(6)\,(2)^3$	.	9979200	.	4989600	.	.	.	.	4989600	.	1663200
$(6)\,(3)\,(1)^3$	.	.	−13305600	−13305600	−26611200	−13305600	−8870400	−26611200	−13305600	−13305600	−4435200
$(6)\,(3)\,(2)\,(1)$	.	.	13305600	13305600	.	13305600	26611200	.	13305600	13305600	13305600
$(6)\,(3)^2$	.	.	.	.	.	.	−8870400	.	.	.	−4435200
$(6)\,(4)\,(1)^2$	.	.	.	9979200	.	.	.	19958400	9979200	19958400	9979200
$(6)\,(4)\,(2)$	.	.	.	−9979200	.	.	.	.	−9979200	.	−9979200
$(6)\,(5)\,(1)$	.	.	.	.	.	.	.	.	.	−15966720	−15966720
$(6)^2$	.	.	.	.	.	.	.	.	.	.	6652800
$(7)\,(1)^5$	34214400	17107200	11404800	2851200	34214400	17107200	5702400	11404800	5702400	2851200	570240
$(7)\,(2)\,(1)^3$	−34214400	−34214400	−34214400	−17107200	−34214400	−34214400	−22809600	−34214400	−22809600	−17107200	−5702400
$(7)\,(2)^2\,(1)$	.	17107200	.	8553600	.	17107200	17107200	.	17107200	8553600	8553600
$(7)\,(3)\,(1)^3$	.	.	22809600	22809600	.	.	11404800	22809600	11404800	22809600	11404800
$(7)\,(3)\,(2)$	.	.	.	.	.	.	−11404800	.	−11404800	.	−11404800
$(7)\,(4)\,(1)$	.	.	.	−17107200	.	.	.	.	.	−17107200	−17107200
$(7)\,(5)$	.	.	.	.	.	.	.	.	.	.	13685760
$(8)\,(1)^4$	−29937600	−14968800	−9979200	−2494800	−59875200	−29937600	−9979200	−29937600	−14968800	−9979200	−2494800
$(8)\,(2)\,(1)^2$	29937600	29937600	29937600	14968800	.	29937600	29937600	29937600	29937600	29937600	14968800
	16										
$(8)\,(2)^2$	.	−14968800	.	−7484400	.	.	.	.	−14968800	.	−7484400
	32	−32									
$(8)\,(3)\,(1)$	.	.	−19958400	−19958400	.	.	−19958400	.	.	−19958400	−19958400
	24	.	−24								
$(8)\,(4)$	.	.	.	14968800	.	.	.	.	.	.	14968800
	32	−16	−32	32							
$(9)\,(1)^3$	.	.	.	.	53223400	26611200	8870400	53223400	26611200	26611200	8870400
	.	.	.	.	9						
$(9)\,(2)\,(1)$	.	.	.	.	.	−26611200	−26611200	.	−26611200	−26611200	−26611200
	18	.	.	.	9	−18					
$(9)\,(3)$	.	.	.	.	.	.	17740800	.	.	.	17740800
	27	.	−27	.	9	−27	27				
$(10)\,(1)^2$	.	.	.	.	.	.	.	−47900160	−23950080	−47900160	−23950080
	10	.	.	.	10	.	.	−10			
$(10)\,(2)$	.	.	.	.	.	.	.	.	23950080	.	23950080
	30	−20	.	.	10	−20	.	−10	20		
$(11)\,(1)$	.	.	.	.	.	.	.	.	.	43545600	43545600
	22	.	−11	.	11	−11	.	−11	.	11	
(12)	.	.	.	.	.	.	.	.	.	.	−39916800
	36	−12	−24	12	12	−24	12	−12	12	12	−12

CONTROL CHARTS WITH WARNING LINES

By E. S. PAGE

*Statistical Laboratory, University of Cambridge**

Process inspection schemes using control charts with warning lines are considered and the properties of some schemes based on the observations from the last few samples are evaluated. Tables of schemes for controlling the mean of a normal population are given. A method is suggested for controlling both the mean and standard deviation of a population on a single chart; examples are given when the population distribution is normal.

1. Introduction

A problem that arises in industrial applications of statistics is to detect changes in the parameters specifying the quality of the output from a continuous production process, so that some rectifying action can be taken to restore the parameters to satisfactory values. A widely used scheme for this purpose (a *process inspection scheme*) is based on a control chart (Shewhart, 1931). Samples of fixed size are taken at regular intervals and a statistic of the sample (for example, the mean, the range, or the number of defectives) is plotted on the chart; if the sample point falls outside control limit(s) drawn on the chart, i.e. if the statistic differs by more than a given amount from its satisfactory value, rectifying action is taken. This scheme will be referred to as a *single-sample* scheme, since the decision whether or not to take action is based upon a single point on the chart; also in what follows the lines on the chart denoting a serious departure of the statistic from its satisfactory value will be called *action lines* instead of control limits. With any process inspection scheme there will in general be a delay between the point at which the parameters changed and that at which the scheme demands action. During this delay the production fails to meet the quality requirements. On the other hand, in practical cases, process inspection schemes, in particular single-sample schemes, cause action to be taken even when the parameters constantly maintain satisfactory values. In the one case there is a loss due to the production of substandard material, and in the other due to unnecessary interference with the production process. These losses, as a function of the parameters, may be used as a basis for the selection of process inspection schemes or for the comparison of two schemes. When the fraction of the production that is inspected is constant, a convenient assessment of these losses can be obtained from the average number of articles that are inspected before the scheme requires rectifying action when the parameters remain constant; this number, a function of the parameter values, has been called the *average run length* (Page, 1954*a*; cf. Aroian & Levene, 1950). We shall base the choice of inspection schemes on the behaviour of their average run length functions.

Consider a single-sample control chart scheme for controlling the mean of a normal distribution. Suppose that the standard deviation, σ, of the process is known, and that the 'ideal' value for the process mean is μ. Then samples of size N are examined at regular intervals, and if the mean, $\bar{x}$, falls outside the action lines drawn at $\mu \pm B_1\sigma/\sqrt{N}$, rectifying action is taken. The current practice is to choose N, usually small, and to use 'three-sigma'

* Present address: University of Durham.

limits, i.e. to take $B_1 = 3$. These action lines are chosen so that the chance is about 0·002 that a given sample point will cause action to be taken when the mean is at its satisfactory value μ. This consideration gives no guidance about the best sample size to choose nor any information about the average amount of time that will elapse before a change in the mean is noticed. If the fraction of production that may be inspected is fixed, for example, by a limitation on the man-hours for inspection, it is more realistic to choose the scheme with the most suitable average run length function. Tables giving values of N and B_1 have been constructed to enable this choice to be made easily for controlling the mean of a normal distribution (Page, 1954*b*), and similar tables could be computed for other situations. It turns out that the best sample size to use is often much larger than those customarily taken in industrial practice, so that samples should be taken less frequently. But although the theoretically best sample size is large it may not be practically convenient to take samples of this size; further, the quality control engineer has got used to small samples and he may be reluctant to change his habits radically. There is, too, his (correct) feeling that a small sample will spot a very serious change in the mean, and for this reason he will wish to continue taking small samples frequently. It is therefore of interest to seek schemes of the control chart type that are easy to apply, that require only small samples to be taken, and yet retain the advantages of the single-sample schemes using large samples. One method is suggested by a modification of the single-sample scheme that has occasionally been used (Dudding & Jennett, 1944); in this a cluster of 'moderately' extreme sample points is treated as a single point outside the action lines and accordingly action is taken. A point is adjudged 'moderately' extreme if it lies between *warning lines* and the action lines.* Thus for controlling the mean of a normal population with known standard deviation, σ, at an ideal value μ warning lines would be drawn at $\mu \pm B_2\sigma/\sqrt{N}$, where $B_2 < B_1$.

We shall consider rules of the following type:

I. *Choose k, n, N. Take samples of size N. Take action if any point falls outside the action lines or if any k out of the last n points fall outside the warning lines.*

In the next section we show that in a certain sense it is reasonable to consider only two special cases of rule I, and in the following section their properties are evaluated.

2. Restriction of the rules

In the rules of type I only the region of the chart in which a sample point falls is taken into account when deciding whether or not to take action; the position of the sample point within, say, the warning region is not considered. For the rules it is only necessary to count the number of points in the various regions; however, for mathematical convenience it is useful to suppose that a score is assigned to each point according to the region in which it falls. A point outside the warning or action lines is obtained when the quality of the sample falls below the required level; if a mark or score is to be assigned to the sample such a defection will merit some penalty. On the other hand, a point within the warning lines will receive a bonus score. It is reasonable to base a process inspection scheme on these scores; an accumulation of penalties will be taken to indicate a deterioration in quality and suitable rectifying action will be taken. After each sample is taken the total penalty score over the last 'few' samples is examined and action is taken if it is 'large'. In order to make the scheme precise, let a score x_i be assigned to the ith sample, where $x_i = -a, b, c$ $(a > 0, c > b > 0)$,

* A chart using only warning lines has been considered by Weiler (1953).

according as the sample point falls within the warning lines, between the warning and action lines, or outside the action lines, respectively. The scheme is:

II. *Take action after the n-th sample if any of the inequalities*

$$\sum_{i=0}^{r} x_{n-i} > h_{r+1} \quad (r = 0, 1, \ldots, s)$$

are satisfied, where s and the h_r are suitably chosen constants.

We consider the case where s is limited only by the number of samples drawn since action was last taken, so that all the partial sums, working back from the last sample, are examined; and where $h_r = h$, all r, so that the total scores in the last one, two, three, ..., samples are tested to see that none is greater than h. In an earlier paper (Page, 1954*a*) it was shown that this scheme is equivalent to a sequence of linear sequential tests (Barnard, 1946; Wald, 1947) with initial score on the acceptance boundary; if the test ends on the acceptance boundary the test is reapplied, while if it ends on the rejection boundary action is taken. Consequently we consider what restrictions must be placed on k and n in order that the scheme I shall be equivalent to the repeated application of a linear sequential test with boundaries at $(0, h)$ and initial score zero.*

First, we have that an initial sequence of $(k-1)$ points between warning and action lines has total penalty score $(k-1)b$; since this sequence does not require action to be taken

$$(k-1)b < h. \tag{1}$$

If now this sequence is followed by $(n-k)$ points within the warning lines each receiving a bonus score $-a$, and then by another point between the warning and action lines, scheme I requires action to be taken. Hence the total penalty score must be at least h, i.e.

$$kb - (n-k)a \geqslant h. \tag{2}$$

Further, any sequence of $(n-k+1)$ bonus points (i.e. between the warning lines) is equivalent to restarting the scheme; in particular, the set consisting of $(k-1)$ points between the warning and action lines followed by $(n-k+1)$ bonus points must have a total penalty score at most zero. Hence

$$(k-1)b - (n-k+1)a \leqslant 0; \tag{3}$$

indeed, if this inequality were not satisfied a finite number of such sets of n points would cause the total penalty score to exceed h and action to be taken, contrary to the conditions I. The combination of (1) and (2) gives

$$b/a > n-k, \tag{4}$$

and hence with (3) we obtain

$$(n-k) < (n-k+1)/(k-1), \tag{5}$$

since $k \geqslant 2$ for the warning lines to be distinct from the action lines. It follows that

$$n < \frac{(k-1)^2}{k-2} = k + \frac{1}{k-2}. \tag{6}$$

For scheme I, k and n are integral and $2 \leqslant k \leqslant n$, so that (6) can be satisfied only for

(i) $k = 2$, any n,

or (ii) $k = n$.

* The restriction on the initial score is unnecessary; if the initial score is Z $(0 \leqslant Z < h)$, inequalities are obtained in a similar way and found to imply $Z = 0$.

The two rules we consider therefore are:

III. *Choose n, N. Take samples of size N. Take action if two points out of any sequence of n fall between the warning and action lines or if any point falls outside the action lines.*

IV. *Choose n, N. Take samples of size N. Take action if n consecutive points fall between the warning and action lines or if any point falls outside the action lines.*

In these two cases it cannot happen that some sets of points not requiring action to be taken have greater total penalty scores than some sets leading to action; for any other values of k, n the anomalous position can, however, occur. The above theory admits the corollary that no other types of warning lines, more or less extreme than those considered above, may be introduced and yet permit the equivalence of a rule of type I based on the last n points, and a rule of type II. In the next section we derive the average run lengths of the schemes III and IV.

3. The average run lengths of the rules

Although we shall not use the following method, for completeness we remark that the average run length of the general rule of type I may be evaluated by enumerating the possible combinations of the last $(n-1)$ points on the chart such that action has not been required, and treating the combinations as the states of a discrete Markov chain (e.g. Bartlett, 1953; Feller, 1950). The matrix of transition probabilities, $\mathbf{P}$, may be written down: the vector giving the probability that the chart is in a given state after the rth point is $\mathbf{p}_r = \mathbf{P}^r\mathbf{p}_0$, where $\mathbf{p}_0$ is the vector specifying the initial state. Accordingly, the probability that action has not been taken up to and including the rth sample is the sum of the elements of $\mathbf{p}_r$, i.e. $\mathbf{l}'\mathbf{p}_r$, where $\mathbf{l}' = (1, 1, \ldots, 1)$. Hence the probability that action is taken immediately after the rth sample is $\mathbf{l}'(\mathbf{p}_{r-1} - \mathbf{p}_r)$, so that the average run length is given by

$$L = \mathbf{l}' \sum_{r=1}^{\infty} r(\mathbf{p}_{r-1} - \mathbf{p}_r) N,$$

i.e.

$$L = \mathbf{l}'(\mathbf{I} - \mathbf{P})^{-1}\mathbf{p}_0 N, \tag{7}$$

where N is the size of the sample.

In the general case the size of the matrix $\mathbf{P}$ increases rapidly with n so that the labour involved in inverting $\mathbf{I} - \mathbf{P}$ is considerable. Fortunately, the transition matrices for the rules III and IV are of a simple type and the states can be more simply specified so that the average run lengths can be found with little trouble. Let the probabilities that a given point falls between the warning lines, between the warning and action lines, and outside the action lines be p_0, p_1, p_2 respectively. Suppose that a suitable convention is made for points falling on the lines; for example, that any point falling on a line is to be regarded as falling into the adjacent more extreme region of the chart. With such a convention, clearly $p_0 + p_1 + p_2 = 1$.

Consider first rule III; it has been shown that this rule is equivalent to a sequence of linear sequential tests with initial score on the acceptance boundary. By inspection it is seen that appropriate scores to assign to the sample point are given by

$$a = 1, \quad b = n - 1, \quad c = n, \tag{8}$$

and action is to be taken if the total penalty score rises a height $h = n$ above its previous least value; or equivalently if the sequential test ends on the rejection boundary given by

$h = n$. When action is not required the position of the total penalty score relative to its previous minimum or of the cumulative score in the test is specified by one of the integers $0, 1, \ldots, n-1$. These may be regarded as the states of a Markov chain and the transition matrix $\mathbf{P}$ is seen to be

$$\mathbf{P} = \begin{bmatrix} p_0 & 0 & \ldots & \ldots & 0 & p_1 \\ p_0 & 0 & \ldots & \ldots & 0 & 0 \\ 0 & p_0 & 0 & \ldots & 0 & 0 \\ & & & \ldots & & \\ 0 & 0 & \ldots & \ldots & p_0 & 0 \end{bmatrix}. \tag{9}$$

The average run length is then given by equation (7). Alternatively, let L_i be the average number of observations drawn before action is taken when the state is i. Then L_0 is the average run length of rule III. We have, by taking expectations conditional upon the result of the first sample,

$$\begin{aligned} L_0 &= p_0(N + L_0) + p_1(N + L_{n-1}) \\ &= N + p_0 L_0 + p_1 L_{n-1}. \end{aligned} \tag{10}$$

Similarly, we obtain the set of equations

$$\left.\begin{aligned} L_0 &= N + p_0 L_0 + p_1 L_{n-1} \\ L_1 &= N + p_0 L_0 \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ L_i &= N + p_0 L_{i-1} \end{aligned}\right\} \quad (1 \leqslant i < n-1). \tag{11}$$

The last $n-1$ of these equations may be solved successively for $L_1, \ldots, L_{n-1}$ in terms of L_0. and the first equation used to evaluate L_0. We obtain the average run length of rule III:

$$L = \frac{(1 - p_0 + p_1 - p_1 p_0^{n-1})N}{(1 - p_0)(1 - p_0 - p_1 p_0^{n-1})} \quad (n \geqslant 2). \tag{12}$$

Scores for the various positions of the sample point on the chart for the repeated sequential tests equivalent to rule IV can be taken to be $a = n-1$, $b = 1$, $c = n$ and $h = n$. There are again only n possible states before action is taken for the Markov chain. The average run length is L_0, where L_0 is given by the equations

$$\left.\begin{aligned} L_i &= N + p_0 L_0 + p_1 L_{i+1} \\ L_{k-1} &= N + p_0 L_0 \end{aligned}\right\} \quad (0 \leqslant i < k-1). \tag{13}$$

We obtain the average run length of rule IV:

$$L = \frac{(1 - p_1^n)N}{1 - p_0 - p_1 + p_0 p_1^n}. \tag{14}$$

4. Choice of scheme

In each of the two types of rule, III and IV, there are certain constants that may be chosen to give the scheme adopted some desired properties. Consider a two-sided chart with two sets of warning and action lines symmetrically placed about the ideal value for the parameter, or a one-sided chart with one warning and one action line. The disposable constants are then the sample size, N, the number of sample points considered, n, and the positions of the warning and action lines; the scheme may also be of one of two types. This may be contrasted with the single-sample control chart scheme in which there are only two disposable

constants, the sample size and the position of the action lines. In the single-sample case the scheme could be chosen to give a specified average run length for some unsatisfactory value of the parameter and maximum average run length for the ideal value (Page, 1954*b*); that is to say, the amount of scrap at a certain value of the parameter that can be tolerated is stated and the scheme is chosen to satisfy this requirement and to give the longest run without action being taken when the quality is satisfactory. A corresponding choice for a scheme with warning lines could be made by selecting two unsatisfactory values of the parameter and the average run lengths that can be tolerated at these values and then choosing the disposable constants to gain maximum average run length on the ideal quality; in this way the average run length will have desirable properties over a range of the parameter. In order to illustrate this we consider schemes for controlling the mean of a normal population with known variance, σ^2, for two-sided deviations from the ideal mean. Suppose that it is desired to obtain a scheme with average run lengths of 100 and 25 when there are departures from the ideal mean of amounts $0{\cdot}4\sigma$ and $0{\cdot}8\sigma$ respectively. The large number of disposable constants would make it laborious to find the scheme of the types considered which has the greatest run length on ideal quality, but one scheme with approximately the run lengths stated above is that of type IV with $k = 4$, sample size $N = 20$, warning lines at $\mu \pm 1{\cdot}5\sigma/\sqrt{N}$, and action lines at $\mu \pm 2{\cdot}875\sigma/\sqrt{N}$. We can compare this scheme with the best single-sample schemes for detecting the two sizes of departure from the mean. From the tables given in an earlier paper (Page, 1954*b*, Tables 2*a–c*) the single-sample scheme having average run length 100 when the mean is $\mu \pm 0{\cdot}4\sigma$ and maximum average run length when the mean is μ has action lines drawn at $\mu \pm 2{\cdot}82\sigma/\sqrt{N}$, and sample size $N = 70$. Similarly, the single-sample scheme with action lines at $\mu \pm 2{\cdot}83\sigma/\sqrt{N}$ and sample size $N = 17$ has $L = 25$ at mean $\mu \pm 0{\cdot}8\sigma$ and maximum average run length at mean μ. The average run lengths for the three schemes for several values of λ, where the mean is $\mu \pm \lambda\sigma$, are shown in Table 1. It is seen that the rule of type IV has approximately the same run length as the single-sample scheme with sample size 17 for large deviations ($> 0{\cdot}8\sigma$) from the true mean; on the other hand, its run length for smaller deviations ($0{\cdot}2 < |\lambda| < 0{\cdot}4$) is rather less than that of the single-sample scheme. Again it is seen that, while the type IV scheme compares unfavourably with the second single-sample scheme ($N = 70$, $B_1 = 2{\cdot}82$) in its behaviour

Table 1. *Average run lengths of control chart schemes*

λ	Single-sample schemes		Type IV scheme
	$N = 17$ $B_1 = 2{\cdot}83$	$N = 70$ $B_1 = 2{\cdot}82$	$n = k = 3$ $N = 20$ $B_1 = 2{\cdot}875$ $B_2 = 1{\cdot}5$
0·0	3652	14600	3368
0·2	780	560	545
0·4	146	100	99
0·6	47	71	42
0·8	25	70	26
1·0	19	70	21

on good quality, it gains considerably for the large departures ($|\lambda| > 0{\cdot}4$); clearly for a single-sample scheme to cause action to be taken at least one sample point must be plotted. The situation we are considering is where an amount N/f of production is output between visits from the inspector (where f is the fraction of output sampled); thus the relevant quantity to consider for even extreme deviations from the mean is the average run length, or approximately the sample size.

The choice of the best sample size could be examined in the same way as in the earlier paper, but it would be extremely laborious to do so. In general, there will exist a value of N giving maximum average run length when the production has the ideal quality and specified average run lengths at two values of the mean. The effects of different sample sizes are illustrated by examples of rules of type IV which have average run lengths of 100 and 25 at $\mu \pm 0{\cdot}4\sigma, \mu \pm 0{\cdot}8\sigma$ respectively; the sample size, N, the positions of the warning and action lines, B_1 and B_2, the number, n, of consecutive samples considered, and the average run length, L, on the ideal mean are shown for five schemes in Table 2. There is, of course, an upper limit to the size of the sample that can be used for the scheme to have the desired run lengths. It is, however, again preferable to use larger samples if this is practically possible.

Table 2. *Average run lengths of type IV schemes*

N	n	B_1	B_2	$L(0{\cdot}8\sigma)$	$L(0{\cdot}4\sigma)$	$L(0)$
5	3	3·25	1·25	25·3	118·0	584·6
10	3	3·125	1·25	25·2	103·9	1096·9
15	3	3·00	1·375	25·6	104·2	2286·4
20	3	2·75	1·75	25·0	101·6	3155·9
23	3	2·625	2·0	25·9	92·3	2607·8

5. THE TABLES

In the Appendix are given tables of the average run lengths of schemes for controlling the mean of a normal population with known standard deviation. The run lengths are tabulated for some schemes of type IV with sample sizes 5, 10, 15 and 20, taking into account the means of the last three or four samples drawn ($n = 3$ or $n = 4$). In order to use the tables to determine the suitable control chart scheme to be used it is necessary to decide first upon the average run lengths that can be permitted at two process means different from the ideal one, and which of the above sample sizes it is most convenient to use. A scheme approximately satisfying these requirements may then be selected by inspection of the tables, possibly using some rough interpolation. The schemes shown in Tables 1 and 2 serve as examples of the method.

It has been remarked that the use of moderately sized samples improves the average run-length function; consequently such samples are to be preferred where possible. For samples of ten and over the average run length for departures from the ideal mean of more than a standard deviation is near the sample size; this is easily seen since the probability that a point falls outside the action lines is nearly unity.

Similar tables could be constructed for the average run lengths of schemes of type III. The two types of scheme are, of course, identical when the number of sample points

considered is two. A few calculations indicate that the two types of scheme have very similar properties and that, corresponding to a given scheme of one type, there exist schemes of the other type with approximately the same average run lengths.

6. Simultaneous control of the mean and standard deviation

It is often required to control simultaneously both the mean and the standard deviation of a normal population. For example, when goods are being manufactured to a specification a change in either the process mean or standard deviation would alter the fraction of defective articles produced; one process inspection scheme for this situation (Jennett & Welch, 1939) is based upon the ratio $(U-\bar{x})/s$, where U is an upper tolerance limit, and $\bar{x}$, s^2 sample estimates of mean and variance. Another possibility is to use the easily calculated estimate of the standard deviation from the sample range, w, in place of s and so base the scheme upon $(U-\bar{x})/w$. The more usual procedure in practice is to keep two charts, one recording the means and the other the ranges, of the samples. Before setting up any of these schemes it is important to consider both the type of rectifying action that will be taken and the changes that are likely to occur. If changes in the mean and standard deviation require different rectifying actions it is necessary that the inspection scheme indicate the type of change and not merely its existence. Again, the choice of scheme will be influenced by the relative frequency of occurrence and importance of the types of change. For example, if changes in the standard deviation are very infrequent there is little point in estimating σ from each sample as would be necessary in Jennett & Welch's rule. The same considerations will hold to decide the average run lengths for various types of change that the scheme must achieve. These remarks lead us to consider the possibility of controlling both mean and standard deviation on a single chart for means when changes in the standard deviation are relatively rare or unimportant. The remainder of this paper is devoted to the development of such a scheme and to comparing it with a conventional scheme.

The control chart scheme using charts for the mean and range are often of the conventional type with action lines drawn so that the probability that a sample point falls outside the action lines on a given chart when the parameters have their ideal values is an assigned amount, α ($\alpha \doteqdot 1/500$ gives the conventional 3-sigma limits on the mean chart). Alternatively and preferably, the action limits could be determined for each chart separately on the basis of the average run length as described above and elsewhere (Page, 1954*b*), so that the losses incurred by the scheme may be estimated. This approach, of course, will not be very accurate for some values of the parameters since, for example, a change in the standard deviation affects the run length of both charts. Consider a single-sample scheme for both charts so that action is taken if the mean of any sample, $\bar{x}$, falls outside the range $(\mu - B\sigma/\sqrt{N}, \mu + B\sigma/\sqrt{N})$, or if the range w is greater than $v_1\sigma$, where v_1 is suitably chosen. Since the mean and range of samples from a normal population are independently distributed with frequency elements $f(\bar{x} \mid \mu', \sigma')\,d\bar{x}$, $g(w \mid \sigma')\,dw$, where μ', σ' are the process mean and standard deviation respectively, the probability that a given sample does not cause action to be taken is q_1', q_2', where

$$q_1' = \int_{\mu - B\sigma/\sqrt{N}}^{\mu + B\sigma/\sqrt{N}} f(\bar{x} \mid \mu', \sigma')\,d\bar{x} = 1 - p_1', \tag{15}$$

$$q_2' = \int_0^{v_1\sigma} g(w \mid \sigma')\,dw = 1 - p_2. \tag{16}$$

Clearly q_1', q_2' are the probabilities that the mean and range respectively fall within the action lines. The run length of the charts together is distributed as the smaller of two independent geometric variates with parameters p_1', p_2'; accordingly, the combined run length, l, is a geometric variable with parameter $1-q_1'q_2'$,

$$P\{l=rN\} = (q_1'q_2')^{r-1}(1-q_1'q_2'). \tag{17}$$

The average run length is thus

$$L = \frac{N}{1-q_1'q_2'}, \tag{18}$$

or, in terms of the average run length of the charts separately, is

$$L = \frac{L_1L_2}{L_1+L_2-1}, \tag{19}$$

with an obvious notation. A numerical example is given in the next section.

We turn now to investigate what control can be achieved by the use of the mean chart only, thus avoiding the need to calculate and plot ranges. A possible scheme makes use of the warning lines in another way. Consider the rule:

V. *Choose n, N. Plot the means of samples of size N on a chart on which are drawn two warning and two action lines. Take action if*:

(i) *any point falls outside the action lines.*

or (ii) *n consecutive points fall outside the warning lines,*

or (iii) *two out of any set of n consecutive points fall outside opposite warning lines.*

A sequence of n points outside a warning line is evidence that the process mean has moved in the corresponding direction; similarly, the occurrence of the means of two near samples outside opposite warning lines points to an increase in the spread of the distribution. This scheme differs from one based on the range in that it depends on the variation between samples and not on that within samples. Such a scheme cannot be expected to be very sensitive in detecting increases in the standard deviation and will become less sensitive as the sample size increases: however, it may serve to keep a check on the standard deviation when control of the mean is of prime importance. In order to evaluate the average run length, L_0, of rule V let $L_i'(L_i'')$ $i = 1, \ldots, n-1$, be the average further number of articles drawn before action is taken when the last i sample points have fallen between the upper (lower) warning and action lines.

Let the probabilities that a sample point falls between the upper (lower) warning and action lines be $r(s)$. Then in the notation of § 3 we have $r+s = p_1$. By considering expectations conditional upon the result of the first sample we have

$$L_0 = p_0(L_0+N)+r(L_1'+N)+s(L_1''+N)+p_2N.$$

Therefore

$$L_0 = p_0L_0+rL_1'+sL_1''+N. \tag{20}$$

Again, by considering expectations conditional on the result of the next sample when the last i points have fallen between the upper warning and action lines, we obtain

$$L_i' = p_0L_0+rL_{i+1}'+N \quad (i = 1, 2, \ldots, n-2), \tag{21}$$

$$L_{n-1}' = p_0L_0+N. \tag{22}$$

A similar set of equations is obtained for the L_i''. Hence we have

$$L_1' = \frac{(N + p_0 L_0)(1 - r^{n-1})}{1 - r}, \tag{23}$$

and similarly

$$L_1'' = \frac{(N + p_0 L_0)(1 - s^{n-1})}{1 - s}. \tag{24}$$

These values substituted in (20) yield the solution

$$L_0 = \frac{(1 - rs - r^n - s^n + sr^n + rs^n) N}{p_2 + rs(1 + p_0) + p_0(r^n + s^n - sr^n - rs^n)}. \tag{25}$$

A similar rule for simultaneously controlling the mean and standard deviation is provided by rule III without modification. If the two 'warning' points causing action to be taken are on the same side of the ideal mean a shift in the mean in that direction will be suspected, while if they are on opposite sides an increase in the standard deviation will be suspected. Of course with this rule, if a change in the standard deviation occurs and action is taken because of points outside the warning lines, the two points are just as likely to be on the same side of the ideal mean as they are to be on opposite sides. Again, action could be taken after a change in the standard deviation because of a point outside the action lines. Consequently the wrong sort of change is more likely to be suggested by the scheme when a change in the standard deviation happens. This is of little importance if the rectifying action is independent of the suspected type of change. However, the correct inference is more likely to be drawn with rule V than with rule III.

The rules of this section have nothing corresponding to the lower control limit for the range sometimes drawn in the conventional method. The purpose of such a line is to enable the occurrences of samples with small range to be investigated in the hope that it will be possible to reduce the process standard deviation. It would, of course, be possible to introduce a condition calling for investigation if two successive samples both had means differing little from the required process mean; however, we shall not consider the consequences of such a complication.

7. Numerical examples of simultaneous control

Consider schemes to control the mean and standard deviation of a normal population. Suppose that both mean and range charts with action lines only are kept so that the average run length, L, of the scheme is given by equations (18) and (19). If the action lines are chosen so that the process inspection schemes formed by taking each of the charts on its own have equal average run lengths when both the process mean and standard deviation are at their ideal values, then the average run length of the rule using both charts with these parameter values is approximately half that of the separate rules. For different values of the mean while the standard deviation is the same, the run length of the range chart is unchanged; consequently the run length of the combination is a little less than that of the mean chart, and approaches it as the change in the mean increases. On the other hand, for different values of the standard deviation but the same mean, both L_1 and L_2 are changed. The average run lengths for a specific scheme with mean μ', and standard deviation σ' are shown in Tables 3 and 4. Let the action limits be drawn in the 'conventional' positions, i.e. so that the probability that a sample taken from ideal quality gives a point outside the action

line is 1/500 for each chart. Thus the lines on the mean chart are drawn at $\mu \pm 3{\cdot}09\sigma/\sqrt{N}$, and for samples of $N = 5$ the range chart has a line at $w_1 = 5{\cdot}24\sigma$. We then have

$$q_1 = \int_{(-B-\lambda\sqrt{N})/K}^{(B-\lambda\sqrt{N})/K} \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}t^2}\, dt, \tag{26}$$

$$q_2 = \int_0^{w_1/K\sigma} g(w)\, dw. \tag{27}$$

For comparison, the average run lengths of one of the combined rules are shown in the tables; the scheme chosen has the same sample size, $N = 5$, and $B_1 = 3{\cdot}00$, $B_2 = 2{\cdot}00$, $n = 3$. With these values the combined rule has approximately the same average run length on the ideal quality as the scheme based upon both mean and range charts. It is seen that the rules gives fair control against changes in the standard deviation, and for changes in the mean has a smaller average run length than the rule using the two charts. Consequently when it is necessary only to keep an eye on the standard deviation while controlling the mean of a normal distribution a scheme of type V may be suitable. If so, the labour of keeping both mean and range charts can be avoided.

Table 3. *Average run lengths of schemes*, $N = 5$, $\mu' = \mu + \lambda\sigma$, $\sigma' = \sigma$

$\lvert\lambda\rvert$	Rule using			Combined rule
	Mean	Range	Mean and range	
0·0	2500	2500	1250	1383
0·2	1149	2500	789	767
0·4	359	2500	314	257
0·6	125	2500	119	90
0·8	52	2500	51	38
1·0	25	2500	25	20

Table 4. *Average run lengths of schemes*, $N = 5$, $\mu' = \mu$, $\sigma' = K\sigma$

K	Rule using		Combined rule
	Range	Mean and range	
1·0	2500	1250	1383
1·25	198	158	241
1·5	51	38	91
1·75	24	18	50
2·0	15	12	33
2·25	11	9	25
2·5	9	7	20

A short table of schemes of type V is given in the Appendix, Table 2, for controlling a normal population with samples of five measurements, using at most a sequence of three such samples. The average run lengths for changes in the mean are shown in Table 2*a*, and for changes in the standard deviation in Table 2*b*.

I wish to thank Dr D. R. Cox for many helpful discussions on the subject of this paper, and the Director, Mathematical Laboratory, Cambridge, for permission to use the EDSAC for the calculation of the tables.

REFERENCES

AROIAN, L. A. & LEVENE, H. (1950). *J. Amer. Statist. Ass.* **45**, 520.
BARNARD, G. A. (1946). *J.R. Statist. Soc. Suppl.* **8**, 1.
BARTLETT, M. S. (1953). *Proc. Camb. Phil. Soc.* **49**, 263.
DUDDING, B. P. & JENNETT, W. J. (1944). *Quality Control Chart Technique.* London: General Electric.
FELLER, W. (1950). *Probability Theory and its Application.* New York: Wiley.
JENNETT, W. J. & WELCH, B. L. (1939). *J.R. Statist. Soc. Suppl.* **6**, 80.
PAGE, E. S. (1954*a*). *Biometrika*, **41**, 100.
PAGE, E. S. (1954*b*). *J.R. Statist. Soc.* B, **16**, 131.
SHEWHART, W. A. (1931). *Economic Control of Quality of Manufactured Product.* New York: Macmillan.
WALD, A. (1947). *Sequential Analysis.* New York: Wiley.
WEILER, H. (1953). *J. Amer. Statist. Ass.* **48**, 816.

APPENDIX

Tables of average run lengths of rules IV

Charts with action lines at $\mu \pm B_1\sigma/\sqrt{N}$, warning lines at $\mu \pm B_2\sigma/\sqrt{N}$.

Table 1*a*. $N = 5$

		$n=3$					$n=4$					
	λ \ B_2	1·00	1·25	1·50	1·75	2·00	1·00	1·25	1·50	1·75	2·00	B_2 / λ
$B_1 =$ 3·00	0·0	202	503	1074	1598	1802	527	1190	1692	1830	1850	0·0
	0·2	137	285	527	751	855	305	583	798	872	887	0·2
	0·4	63	102	160	221	261	112	175	236	269	280	0·4
	0·6	32	43	57	75	90	47	62	79	93	100	0·6
	0·8	19	23	27	32	38	25	29	34	39	42	0·8
	1·0	14	15	16	18	20	16	17	19	20	21	1·0
	1·2	11	11	12	13	13	12	12	12	13	13	1·2
	1·4	9	9	9	9	9	8	8	8	9	9	1·4
	1·6	7	7	7	7	7	7	7	7	7	7	1·6
	1·8	6	6	6	6	6	6	6	6	6	6	1·8
$B_1 =$ 3·125	0·0	208	549	1326	2251	2691	579	1513	2451	2758	2806	0·0
	0·2	142	312	637	1008	1213	337	723	1100	1251	1283	0·2
	0·4	66	111	184	274	344	122	206	300	360	382	0·4
	0·6	33	45	63	87	110	51	70	94	115	128	0·6
	0·8	20	24	29	36	44	27	32	39	45	51	0·8
	1·0	14	16	17	19	22	18	19	21	23	25	1·0
	1·2	11	12	12	13	13	13	13	13	14	14	1·2
	1·4	9	9	9	9	9	9	9	9	10	10	1·4
	1·6	7	7	7	7	7	7	7	7	7	7	1·6
	1·8	6	6	6	6	6	6	6	6	6	6	1·8
$B_1 =$ 3·25	0·0	213	585	1575	3110	4041	621	1854	3518	4202	4319	0·0
	0·2	146	334	747	1331	1726	364	873	1504	1810	1881	0·2
	0·4	68	118	207	335	452	132	237	378	483	527	0·4
	0·6	35	48	69	100	134	54	78	110	143	165	0·6
	0·8	21	25	31	40	50	29	35	43	53	62	0·8
	1·0	15	16	18	21	24	19	21	23	25	28	1·0
	1·2	12	12	13	13	14	14	14	15	15	16	1·2
	1·4	9	9	10	10	10	10	10	10	10	11	1·4
	1·6	8	8	8	8	8	8	8	8	8	8	1·6
	1·8	6	6	6	6	6	6	6	6	6	6	1·8

Table 1*b*. $N = 10$

		$n=3$					$n=4$					
	λ \ B_2	1·00	1·25	1·50	1·75	2·00	1·00	1·25	1·50	1·75	2·00	B_2 / λ
$B_1 =$ 2·875	0·0	387	896	1684	2249	2434	933	816	2333	2456	2473	0·0
	0·2	191	336	531	689	763	360	572	720	774	786	0·2
	0·4	70	90	119	149	171	98	127	156	175	183	0·4
	0·6	34	38	43	49	55	42	46	51	56	59	0·6
	0·8	22	22	23	24	26	24	25	25	26	27	0·8
	1·0	15	16	16	16	16	16	16	16	16	16	1·0
$B_1 =$ 3·00	0·0	404	1005	2149	3197	3604	1055	2380	3384	3659	3700	0·0
	0·2	201	374	643	904	1044	405	709	963	1069	1095	0·2
	0·4	71	97	134	177	213	107	146	188	220	236	0·4
	0·6	35	40	47	55	63	45	51	58	65	70	0·6
	0·8	23	24	25	26	28	26	27	28	29	30	0·8
	1·0	16	16	17	17	17	17	17	17	17	18	1·0
$B_1 =$ 3·125	0·0	417	1097	2651	4502	5382	1159	3026	4901	5516	5613	0·0
	0·2	209	408	761	1174	1436	447	862	1285	1491	1545	0·2
	0·4	74	104	149	208	264	116	165	225	277	306	0·4
	0·6	37	43	51	61	72	48	56	65	75	84	0·6
	0·8	24	25	27	29	31	28	29	31	32	34	0·8
	1·0	17	17	18	18	18	18	19	19	19	19	1·0
$B_1 =$ 3·25	0·0	426	1169	3150	6220	8081	1241	3707	7036	8404	8637	0·0
	0·2	216	437	879	1502	1960	483	1022	1700	2094	2208	0·2
	0·4	76	110	164	241	325	124	184	266	348	399	0·4
	0·6	39	45	54	67	82	51	60	73	87	100	0·6
	0·8	25	27	28	31	34	30	32	33	36	38	0·8
	1·0	18	19	19	19	20	20	20	20	20	21	1·0

APPENDIX (*cont.*)

Table 1*c*. $N = 15$

		$n=3$					$n=4$					
	λ \ B_2	1·00	1·25	1·50	1·75	2·00	1·00	1·25	1·50	1·75	2·00	B_2 \ λ
$B_1 =$ 2·75	0·0	548	1163	1929	2367	2492	1201	2034	2423	2505	2516	0·0
	0·2	208	324	457	556	601	344	483	573	607	615	0·2
	0·4	67	80	95	110	123	86	100	114	124	129	0·4
	0·6	35	36	38	40	42	39	40	42	43	44	0·6
	0·8	23	23	23	23	23	23	23	23	23	24	0·8
	1·0	17	17	17	17	17	17	17	17	17	17	1·0
$B_1 =$ 2·875	0·0	581	1344	2526	3373	3651	1399	2724	3500	3684	3710	0·0
	0·2	222	365	552	714	800	392	594	748	814	831	0·2
	0·4	71	86	106	128	147	94	114	134	149	158	0·4
	0·6	37	39	41	44	47	42	44	46	48	50	0·6
	0·8	24	24	24	24	25	25	25	25	25	25	0·8
	1·0	18	18	18	18	18	18	18	18	18	18	1·0
$B_1 =$ 3·00	0·0	607	1508	3223	4795	5406	1582	3570	5076	5489	5549	0·0
	0·2	234	403	656	914	1071	439	722	977	1101	1137	0·2
	0·4	74	93	117	147	175	103	127	156	181	196	0·4
	0·6	39	41	44	48	53	46	48	51	54	57	0·6
	0·8	25	25	26	26	27	27	27	27	27	27	0·8
	1·0	18	18	18	18	18	18	18	18	18	18	1·0
$B_1 =$ 3·125	0·0	626	1645	3977	6753	8073	1738	4539	7351	8275	8420	0·0
	0·2	244	438	765	1159	1439	482	862	1270	1502	1574	0·2
	0·4	78	98	128	167	208	110	141	180	218	244	0·4
	0·6	41	44	48	53	58	49	52	56	61	65	0·6
	0·8	27	27	27	28	29	29	29	29	30	30	0·8
	1·0	19	19	19	19	19	19	19	19	19	19	1·0

Table 1*d*. $N = 20$

		$n=3$					$n=4$					
	λ \ B_2	1·00	1·25	1·50	1·75	2·00	1·00	1·25	1·50	1·75	2·00	B_2 \ λ
$B_1 =$ 2·75	0·0	730	1550	2571	3156	3323	1601	2712	3230	3340	3355	0·0
	0·2	226	333	458	558	608	355	484	575	614	625	0·2
	0·4	70	79	90	102	111	86	95	105	112	117	0·4
	0·6	37	38	39	40	41	40	40	41	41	42	0·6
	0·8	25	25	25	25	25	25	25	25	25	25	0·8
	1·0	21	21	21	21	21	21	21	21	21	21	1·0
$B_1 =$ 2·875	0·0	774	1792	3368	4497	4868	1866	3632	4667	4913	4947	0·0
	0·2	241	372	545	703	795	402	586	737	810	832	0·2
	0·4	74	85	99	115	129	93	106	120	132	140	0·4
	0·6	39	40	41	43	45	43	44	45	46	46	0·6
	0·8	26	26	26	26	26	26	26	26	26	26	0·8
	1·0	21	21	21	21	21	21	21	21	21	21	1·0
$B_1 =$ 3·00	0·0	809	2010	4297	6394	7208	2110	4761	6768	7318	7399	0·0
	0·2	254	409	638	882	1045	447	701	943	1077	1121	0·2
	0·4	78	91	108	129	150	101	117	136	155	168	0·4
	0·6	42	43	45	47	49	47	48	49	50	52	0·6
	0·8	27	27	27	27	28	28	28	28	28	28	0·8
	1·0	22	22	22	22	22	22	22	22	22	22	1·0
$B_1 =$ 3·125	0·0	834	2194	5302	9003	10764	2317	6053	9802	11033	11226	0·0
	0·2	264	443	735	1095	1376	489	824	1200	1440	1528	0·2
	0·4	81	96	117	144	174	109	129	155	182	203	0·4
	0·6	44	46	48	50	53	50	52	53	55	57	0·6
	0·8	29	29	29	29	29	29	29	29	29	30	0·8
	1·0	22	22	22	22	22	22	22	22	22	22	1·0

APPENDIX (*cont.*)

Tables of average run lengths of rules V

Charts with action lines at $\mu \pm B_1\sigma/\sqrt{N}$, warning lines at $\mu \pm B_2\sigma/\sqrt{N}$. Sample size, $N = 5$.

Table 2*a*. $\mu' = \mu + \lambda\sigma$, $\sigma' = \sigma$.

	B_2 / λ	1·50	1·75	2·00
$B_1 =$ 2·875	0·0	404	722	1021
	0·2	299	457	564
	0·4	128	170	196
	0·6	51	63	73
	0·8	25	29	32
	1·0	15	16	17
$B_1 =$ 3·00	0·0	446	879	1383
	0·2	344	577	767
	0·4	148	211	257
	0·6	57	74	90
	0·8	27	32	38
	1·0	16	18	20
$B_1 =$ 3·125	0·0	481	1033	1830
	0·2	385	713	1037
	0·4	169	260	336
	0·6	63	86	110
	0·8	29	36	44
	1·0	17	19	22
$B_1 =$ 3·25	0·0	507	1173	2340
	0·2	422	856	1384
	0·4	188	314	438
	0·6	69	99	133
	0·8	32	40	50
	1·0	18	21	24

Table 2*b*. $\mu' = \mu$, $\sigma' = K\sigma$.

	B_2 / K	1·50	1·75	2·00
$B_1 =$ 2·875	1·00	404	722	1021
	1·25	117	160	198
	1·50	56	69	79
	1·75	35	41	45
	2·00	26	28	30
	2·25	20	22	23
$B_1 =$ 3·00	1·00	446	879	1383
	1·25	129	186	241
	1·50	61	77	91
	1·75	38	44	50
	2·00	27	31	33
	2·25	22	23	25
$B_1 =$ 3·125	1·00	481	1033	1830
	1·25	140	213	292
	1·50	66	86	105
	1·75	41	48	56
	2·00	29	33	36
	2·25	23	25	27
$B_1 =$ 3·25	1·00	507	1173	2340
	1·25	151	240	349
	1·50	71	95	120
	1·75	44	53	62
	2·00	31	35	39
	2·25	24	26	29

MISCELLANEA

Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution*

By C. W. DUNNETT† and M. SOBEL‡

Cornell University

In a recent paper (Dunnett & Sobel, 1954), a multivariate analogue of Student's t-distribution was defined as the joint distribution of p variates $t_{in} = z_i/s$ $(i = 1, 2, \ldots, p)$. Here the z_i have a non-singular multivariate normal distribution with means 0, common unknown variance σ^2 and known correlation matrix $\{\rho_{ij}\}$ and ns^2/σ^2 has a χ^2_n-distribution, independent of the z_i, with n degrees of freedom. The joint density of the t_{in} is given by

$$f(t_{1n}, \ldots, t_{pn}) = \frac{A^{\frac{1}{2}}\Gamma\{\frac{1}{2}(n+p)\}}{(n\pi)^{\frac{1}{2}p}\Gamma(\frac{1}{2}n)}\left[1+\frac{1}{n}\sum_{i,j} a_{ij}t_{in}t_{jn}\right]^{-\frac{1}{2}(n+p)}, \tag{1}$$

where A is the determinant of the positive definite matrix $\{a_{ij}\} = \{\rho_{ij}\}^{-1}$. In the authors' previous paper expressions and tables for the probability integral and equi-co-ordinate percentage points of (1) were obtained for the bivariate case $(p=2)$. In that paper an equi-co-ordinate P-percentage point was defined as the value of h for which

$$\int_{-\infty}^{h} \cdots \int_{-\infty}^{h} f(t_{1n}, \ldots, t_{pn})\, dt_{1n} \ldots dt_{pn} = P. \tag{2}$$

In this note we shall derive approximations (which are also lower bounds) to the probability integral of (1) applicable in special cases when $p > 2$. These results then can be used to obtain approximations (which are also upper bounds) to any equi-co-ordinate P-percentage point. Equation (11) below shows that the probability integral table (Table 1) of the previous paper can be used when $\rho_{ij} = \frac{1}{2}$ $(i \neq j)$ to obtain numerically the approximations referred to above.

Letting $I = I_n(h_1, h_2, \ldots, h_p)$ denote the left-hand member of (2) with the upper limit of t_{in} replaced by h_i $(i = 1, 2, \ldots, p)$, we have

$$I = \Pr\{t_{1n} < h_1, \ldots, t_{pn} < h_p\} = \Pr\{z_1 < h_1 s, \ldots, z_p < h_p s; n\}. \tag{3}$$

Fixing s as the last variable to be integrated, we can write

$$I = \int_0^\infty G_p(h_1 s, \ldots, h_p s; \{\rho_{ij}\})\, f_n(s)\, ds, \tag{4}$$

where $G_p = G_p(x_1, \ldots, x_p; \{\rho_{ij}\})$ is the c.d.f. of the standardized p-variate normal distribution with correlation matrix $\{\rho_{ij}\}$, and $f_n(s)$ is the probability density function of s with n degrees of freedom.

Assumption 1. The matrix $\{\rho_{ij}\}$ has the structure $\rho_{ij} = b_i b_j$ $(i \neq j)$, where $0 \leqq b_i < 1$ $(i = 1, 2, \ldots, p)$. It follows that $\{\rho_{ij}\}$ is positive definite, since the associated quadratic form $\sum_i (1-b_i^2)x_i^2 + (\sum_i b_i x_i)^2$ is positive for x_i not all zero.

Assumption 2. The upper limits of integration in (3) are non-negative, i.e. $h_i \geqq 0$ $(i = 1, 2, \ldots, p)$.

We note that when $\rho_{ij} = \rho \geqq 0$ $(i \neq j)$, Assumption 1 is satisfied since this occurs when $b_i = \sqrt{\rho}$ $(i = 1, 2, \ldots, p)$.

If we let $y_0, y_1, \ldots, y_p$ denote independent, normally distributed chance variables with zero means and unit variances, and let $c_i = \sqrt{(1-b_i^2)}$, then the joint distribution of the chance variables

$$z_i = c_i y_i - b_i y_0 \quad (i = 1, 2, \ldots, p)$$

is a standardized p-variate normal distribution with correlation matrix $\{\rho_{ij}\}$.

* This research was supported in part by the United States Air Force, through the Office of Scientific Research of the Air Research and Development Command.

† Now at Lederle Laboratories, Pearl River, N.Y.

‡ Now at Bell Telephone Laboratories, Allentown, Pa.

Consider the function G_p in (4) above, hold s fixed, and let $a_i = h_i s$ $(i = 1, 2, \ldots, p)$. Then

$$G_p = \Pr\{c_i y_i - b_i y_0 < a_i \quad (i = 1, 2, \ldots, p)\}$$
$$= \int_{-\infty}^{\infty} \prod_{i=1}^{p} G\left(\frac{a_i + b_i y_0}{c_i}\right) g(y_0)\, dy_0, \tag{5}$$

where $g(y)$ is the standard univariate normal density and $G(y)$ is its c.d.f.

It can easily be shown (see, for example, Kimball, 1951) that for any r non-decreasing, bounded functions $F_i(x)$ $(i = 1, 2, \ldots, r)$ of a chance variable x we have (letting E denote expectation)

$$E\left\{\prod_{i=1}^{r} F_i(x)\right\} \geqq \prod_{i=1}^{r} E\{F_i(x)\}. \tag{6}$$

Applying (6) to (5) gives
$$G_p \geqq \prod_{i=1}^{p} \int_{-\infty}^{\infty} G\left(\frac{a_i + b_i y_0}{c_i}\right) g(y_0)\, dy_0$$
$$= \prod_{i=1}^{p} \Pr\{c_i y_i - b_i y_0 < a_i\} = \prod_{i=1}^{p} G(a_i). \tag{7}$$

Substituting this result in (4) and applying (6) again gives

$$I \geqq \int_0^{\infty} \prod_{i=1}^{p} G(h_i s) f_n(s)\, ds \geqq \prod_{i=1}^{p} \int_0^{\infty} G(h_i s) f_n(s)\, ds$$
$$= \prod_{i=1}^{p} \Pr\{z_i < h_i s;\, n\} = \prod_{i=1}^{p} \Pr\{t_{in} < h_i\}. \tag{8}$$

This lower bound to I does not depend on $\{\rho_{ij}\}$ and is easily calculated from tables of the c.d.f. of the univariate Student t-distribution with n degrees of freedom.

Under Assumptions 1 and 2 a sharper (i.e. higher) lower bound can be obtained by obvious modifications of the above argument. This lower bound depends on the c.d.f. of the bivariate t-distribution considered by Dunnett & Sobel (1954) which is tabulated there for the special case $\rho = \frac{1}{2}$ and $h_1 = h_2 \geqq 0$. The results are for even $p \geqq 2$,

$$I \geqq \prod_{i=1}^{\frac{1}{2}p} \Pr\{t_{2i-1,\,n} < h_{2i-1}, t_{2i,\,n} < h_{2i}\}, \tag{9}$$

and, for odd $p \geqq 3$,
$$I \geqq \Pr\{t_{1n} < h_1\} \prod_{i=1}^{\frac{1}{2}(p-1)} \Pr\{t_{2i,\,n} < h_{2i}, t_{2i+1,\,n} < h_{2i+1}\}. \tag{10}$$

If we replace Assumptions 1 and 2 above by

Assumption 1′. The matrix $\{\rho_{ij}\}$ has the structure $\rho_{ij} = \rho$ $(i \neq j)$, where $0 \leqq \rho < 1$. Clearly, $\{\rho_{ij}\}$ is positive definite.

Assumption 2′. The upper limits of integration in (3) are all equal, i.e. $h_i = h$ $(i = 1, 2, \ldots, p)$; then we can replace (9) and (10) by a single inequality and write for any integer $p \geqq 2$

$$I \geqq [\Pr\{t_{1n} < h, t_{2n} < h\}]^{\frac{1}{2}p}, \tag{11}$$

which is sharper than (10) (when p is odd). The proof of (11) is similar to the above proof, using, instead of (6), the well-known inequality (see, for example, Cramér, 1946, p. 176)

$$\beta_p^{1/p} \geqq \beta_q^{1/q} \quad \text{for} \quad p \geqq q, \tag{12}$$

where β_p denotes the pth absolute moment.

Paulson (1952) suggested an alternative method of obtaining an approximation and lower bound to (3) which is based on the Bonferroni inequality

$$\Pr\{t_{1n} < h_1, \ldots, t_{pn} < h_p\} \geqq 1 - \sum_{i=1}^{p} \Pr\{t_{in} > h_i\}. \tag{13}$$

This method has the advantage that it requires no assumptions. However, when the inequalities (8) through (11) hold, then (8) and hence also (9), (10) and (11) give sharper lower bounds than (13). To show this for (8) let

$$1 - q_i = \Pr\{t_{in} < h_i\}, \tag{14}$$

so that $0 \leqq q_i \leqq 1$. Then we have to prove that

$$\prod_{i=1}^{p} (1 - q_i) \geqq 1 - \sum_{i=1}^{p} q_i. \tag{15}$$

This certainly holds for $p = 1$, and a straightforward mathematical induction shows that (15) holds for all positive integers p.

The approximations (13), (8), (9), (10) and (11) can be used to obtain upper bounds to the equi-co-ordinate percentage points of (1) for $p > 2$. Table 1 compares the approximations with the exact values for the special cases $\rho_{ij} = \frac{1}{2}\ (i \neq j)$; $n = 5, \infty$; $p = 3, 9$ and $P = 0{\cdot}50, 0{\cdot}75, 0{\cdot}95, 0{\cdot}99$; all entries are rounded to the nearest two decimal places. For $n = \infty$ columns (13) and (8) require a table of the univariate normal c.d.f.; columns (10), (11) and the exact values were obtained from unpublished tables of the National Bureau of Standards (1953). For $n = 5$, columns (13) and (8) require a table of the univariate Student c.d.f. with 5 degrees of freedom; columns (10) and (11) and the exact values were computed by numerical integration, i.e. by applying Simpson's rule and using the National Bureau of Standards tables (1953) to evaluate the integral in (4). The exact values for $P = 0{\cdot}95, 0{\cdot}99$ will also appear in another paper by Dunnett where equi-co-ordinate percentage points for the case $\rho_{ij} = \frac{1}{2}\ (i \neq j)$ will be given for $p = 3\,(1)\,9$, $P = 0{\cdot}95, 0{\cdot}99$ and $n = 5\,(1)\,20, 24, 30, 40, 60, 120, \infty$.

Table 1. *Comparison of exact equi-co-ordinate percentage points with approximations for selected values of n, p and P*

P	n	$p=3$					$p=9$				
		Approximations				Exact values	Approximations				Exact values
		(13)	(8)	(10)	(11)		(13)	(8)	(10)	(11)	
0·99	5	4·46	4·46	4·37	4·32	4·21	5·75	5·75	5·61	5·59	5·03
	∞	2·71	2·71	2·70	2·70	2·68	3·06	3·06	3·05	3·05	3·00
0·95	5	2·91	2·90	2·82	2·78	2·68	3·93	3·90	3·79	3·78	3·30
	∞	2·13	2·12	2·10	2·09	2·06	2·54	2·53	2·52	2·51	2·42
0·75	5	1·62	1·55	1·46	1·42	1·32	2·48	2·38	2·28	2·26	1·81
	∞	1·38	1·33	1·28	1·26	1·19	1·91	1·86	1·82	1·82	1·60
0·50	5	1·07	0·89	0·80	0·74	0·62	1·93	1·71	1·60	1·58	1·10
	∞	0·97	0·82	0·74	0·70	0·59	1·59	1·45	1·39	1·38	1·04

For each of the approximations (13), (8), (10) and (11) in Table 1 it is conjectured that further calculation will establish the following properties for the difference $D = D(n, p, P)$ between the approximation and its exact value:

(i) D is increasing with p for each n and P,
(ii) D is decreasing with n for each p and P,
(iii) D is decreasing with P for $n = \infty$ and each p,
(iv) D is parabolic-shaped with P for $n = 5$ and each p.

The values of n, p and P in Table 1 were selected to cover a wide range of practical interest. Since only a limited number of exact values for finite n are known the inequalities considered in this paper, which have a fairly wide application, should prove to be useful.

REFERENCES

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press.

Dunnett, C. W. A multiple comparisons test for comparing several treatments with a control. Unpublished manuscript.

Dunnett, C. W. & Sobel, M. (1954). *Biometrika*, **41**, 153.

Kimball, A. W. (1951). *Ann. Math. Statist.* **22**, 600.

National Bureau of Standards (1953). Personal communication on unpublished tables.

Paulson, E. (1952). *Ann. Math. Statist.* **23**, 239.

Galton's rank-order test

BY J. L. HODGES, JR.

University of California, Berkeley

1. One of the first uses of rank-order in statistics was that of Galton in the study of data referred to him by Charles Darwin (1876). A quantity was measured on each of n treated subjects and also on each of n control subjects, obtaining measurements $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_n$ respectively. The $2n$ measurements were arranged in common increasing order of size, and Galton counted the number, say G, of times that an x of given rank exceeded the y of the same rank. In Galton's case, n was 15 and G was 13, which was regarded as evidence that the first sample came from a population stochastically larger than that from which the second sample came. In modern language, if G is sufficiently large, we reject the null hypothesis that the treatment is without effect, in favour of the alternative that the treatment tends to increase the measurements.*

2. Galton was not able to attach a significance level to his observation, inasmuch as he did not know the distribution of G under the null hypothesis. However, that distribution has recently been discovered in another connexion and proved to be very simple and elegant. Consider the usual penny-tossing game played by Peter and Paul, in which Peter pays one unit to Paul if the penny lands 'heads' and Paul pays one unit to Peter if it lands 'tails'. Suppose we are given that after $2n$ tosses, the contestants are even. Let F denote the number of times that Peter was in the lead (where conventionally we assert that, when the game is even, the player leads who led on the preceding toss). If we identify Peter's winning the kth toss with the event that the kth measurement in the Galton problem is an x, it appears that $F = 2G$.

The conditional distribution of F has been found by Chung & Feller (1949) to be uniform, under the hypothesis that Peter's n victories are randomly distributed among the $2n$ trials. But this is just the distribution of the ranks of the x measurements under the hypothesis that the treatment has no effect. Therefore we may assert, under the null hypothesis, that $P(G \leqq g) = (n-g+1)/n$. For example, to Galton's observation we may attach the significance probability 1/5.

3. The proof of Chung & Feller is by means of a double generating function, and it may be of interest to have an enumerative proof of so simple a result. Consider the class of $\binom{2n}{n}$ possible arrangements of n 0's and n 1's, and for each such arrangement determine its score g by counting those 1's which precede the 0 of the same ordinal number. For example, in the sequence 0 1 1 1 0 0 0 1 the black 1's are counted, since the second 1 precedes the second 0, the third 1 precedes the third 0, but the first and fourth 1's are not counted as they follow the first and fourth 0's respectively. The sequence just given has a total score $g = 2$. For a sequence of length $2n$, the possible values of g are $0, 1, \ldots, n$. We shall denote the score of a sequence $(a_1 \ldots a_{2n})$ by $[a_1 \ldots a_{2n}]$.

THEOREM. The number of arrangements with score x is independent of x.

We shall prove this by defining a mapping T_n which has the following properties:

(*a*) The domain of T_n is the set of all sequences of length $2n$ whose score is positive.

(*b*) The range of T_n is the set of all sequences of length $2n$ whose score is less than n.

(*c*) T_n is a 1-1 function.

(*d*) $[T_n(a_1 \ldots a_{2n})] = [a_1 \ldots a_{2n}] - 1$.

As such a function gives a 1-1 mapping of the arrangements with score x on to those with score $x-1$ for $x = 1, \ldots, n$, it establishes the desired result.

We shall need to consider the first point in the sequence at which there have been equal numbers of 0's and 1's (this corresponds in coin-tossing to the first toss at which the game is even). For each arrangement $(a_1, \ldots, a_{2n})$ let k be the smallest positive integer such that $a_1 + \ldots + a_{2k} = k$. We say that the sequence 'breaks' at $2k$, and note that

$$[a_1 \ldots a_{2n}] = [a_1 \ldots a_{2k}] + [a_{2k+1} \ldots a_{2n}]. \tag{1}$$

* Galton's analysis has been extensively reviewed by Fisher (1945, Ch. III), who points out that Darwin did not actually have two samples of n, but n matched pairs. It may be noted that the pairing did not serve to reduce the variance; in fact, if we test the hypothesis that there is no pair effect we have $F_{14,14} = 0{\cdot}554$. In any case, our present concern is with the problem of two samples of n.

Note that k may equal n. As $2k$ is the *first* equilibrium point, there cannot have been an earlier change of lead, so $[a_1 \dots a_{2k}]$ must be either 0 or k. Further, if $[a_1 \dots a_{2k}] = k$, we must have $a_1 = 1$, $a_{2k} = 0$ and $[a_2 \dots a_{2k-1}] = k-1$.

We now define the functions T_n inductively. Let $T_1(10) = (01)$ and check properties (a)–(d) for this case. Suppose we have defined T_m for $m<n$ satisfactorily. For any $(a_1 \dots a_{2n})$ of positive score, let $T_n(a_1 \dots a_{2n})$ be

$$\text{(i)}\ (a_1 \dots a_{2k} T_{n-k}(a_{2k+1} \dots a_{2n})) \quad \text{if} \quad [a_{2k+1} \dots a_{2n}] > 0,$$

$$\text{(ii)}\ (0 a_{2k+1} \dots a_{2n} 1 a_2 \dots a_{2k-1}) \quad \text{if} \quad [a_{2k+1} \dots a_{2n}] = 0.$$

We must check (a)–(d), of which (a) is obvious. Condition (d) holds for (i) by the induction hypothesis and (1). As for (ii), if $[a_{2k+1} \dots a_{2n}] = 0$, we must have

$$0 < [a_1 \dots a_{2k}] = k = [a_1 \dots a_{2n}], \quad [0 a_{2k+1} \dots a_{2n} 1] = 0, \quad [a_2 \dots a_{2k-1}] = k-1 = [T_n(a_1 \dots a_{2n})].$$

From (d) it appears that the range of T_n is contained in the set of sequences with score less than n; any such sequence has a T_n-inverse under (i) if $[a_{2k+1} \dots a_{2n}] < n-k$, since T_{n-k} satisfies the conditions, and under (ii) otherwise, since when $[a_{2k+1} \dots a_{2n}] = n-k$ we have

$$k > [a_1 \dots a_{2k}] = 0, \quad a_1 = 0, \quad a_{2k} = 1, \quad [a_2 \dots a_{2k-1}] = 0, \quad (a_1 \dots a_{2n}) = T_n(1 a_{2k+1} \dots a_{2n} 0 a_2 \dots a_{2k-1}).$$

The ranges of (i) and (ii) are disjoint and each is invertible, so (c) holds.

4. While the Galton test as defined above applies only when the two samples are of equal size, there is a natural extension to the more general case. For example, let $n_1 = 3$, $n_2 = 11$. The third, sixth and ninth ranking observations in the larger sample divide it into four segments, each containing two observations. If we regard these three observations as 'representing' the larger sample, we may calculate G as before between the two samples of three each, and it is clear that G thus defined has the uniform distribution, since to each arrangement for it there corresponds the same number (27) of arrangements of the original problem. In general, if $n_2 = n_1 + k(n_1+1)$, we may represent the second sample by its observations of rank $k+1$, $2k+2$, ..., $n_1 k + n_1$, and obtain a G which is uniformly distributed over the values $0, 1, \dots, n_1$.

If $n_2 - n_1$ does not happen to be divisible by n_1+1, the above method will not apply exactly, but we may still obtain a uniformly distributed test statistic by randomization. The representative values of the larger sample may be chosen so that the segments into which they partition it differ by at most one in size. If we select one among all such partitionings at random, the resulting G may be shown to be uniformly distributed. Because of the natural repugnance of randomized decisions in such problems, it is probably preferable instead to associate with each arrangement the distribution of G values it has among the partitions, and choose for its definitive G the integer nearest the expectation of this distribution.

An interesting result follows if we let $n_2 \to \infty$. The population from which the second sample was chosen in then known, and we are dealing with the one sample rather than the two sample problem. Our statistic is now definable as the number of the quantities $y_k - F^{-1}(n_1 + 1 - k/n_1 + 1)$ which are positive. That its distribution is still uniform may be seen by a limiting argument from the above, or by considering it as a set of conditioned partial sums.

REFERENCES

CHUNG, KAI LAI & FELLER, W. (1949). On fluctuations in coin-tossing. *Proc. Nat. Acad. Sci., Wash.*, **35**, 605–8.

DARWIN, CHARLES (1876). *The Effects of Cross- and Self-fertilization in the Vegetable Kingdom.* London: John Murray.

FISHER, R. A. (1945). *The Design of Experiments*, 4th ed. Edinburgh: Oliver and Boyd.

On bounds for the normal integral*

BY JOHN T. CHU

Institute of Statistics, University of North Carolina

1. Let $$v = \int_0^x (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}t^2}\,dt \quad (x \geqslant 0). \tag{1}$$

G. Pólya (1949) and J. D. Williams (1946) proved independently that

$$v \leqslant \tfrac{1}{2}(1-e^{-2x^2/\pi})^{\frac{1}{2}}. \tag{2}$$

Two simple questions follow naturally: (i) Is it possible to replace the constant $2/\pi$ in (2) by a smaller quantity without breaking the inequality? (ii) Does there exist a lower bound, in a similar form, for v? We find the following answer.

If for all $x \geqslant 0$, the integral v given by (1) satisfies

$$\tfrac{1}{2}(1-e^{-ax^2})^{\frac{1}{2}} \leqslant v \leqslant \tfrac{1}{2}(1-e^{-bx^2})^{\frac{1}{2}}, \tag{3}$$

then it is necessary and sufficient that $0 \leqslant a \leqslant \frac{1}{2}$ and $b \geqslant 2/\pi$.

The proof of this statement is simple. First,

$$\lim_{x\to 0} v^2/(1-e^{-bx^2}) = (2\pi b)^{-1}.$$

Hence if (3) is true, $b \geqslant 2/\pi$. On the other hand (3) implies $x^2/[-\log(1-4v^2)] \leqslant 1/a$ for all real x. Since the limit of this ratio, as $x \to \infty$, is 2, we have $a \leqslant \frac{1}{2}$. Finally

$$4v^2 = \int_{-x}^{x}\int_{-x}^{x} (2\pi)^{-1} e^{-\frac{1}{2}(s^2+t^2)}\,ds\,dt \geqslant \int_0^{2\pi}\int_0^x (2\pi)^{-1} e^{-\frac{1}{2}r^2}\,r\,dr\,d\theta.$$

Therefore $$v \geqslant \tfrac{1}{2}(1-e^{-\frac{1}{2}x^2})^{\frac{1}{2}}. \tag{4}$$

Pólya showed that as x varies from 0 to ∞, the ratio of the L.H.S. (left-hand side) of (2) to the R.H.S decreases steadily from 1 to a minimum value and then increases steadily. Williams's calculations indicate that, approximately, the minimum value 0·9930 is taken as $x = 1{\cdot}6$. Using a similar method to that of Pólya, it can be shown that the ratio of the L.H.S. of (4) to the R.H.S. is a steadily decreasing function of x for all $x \geqslant 0$; for the derivative of this ratio has the same sign as

$$2(e^{\frac{1}{2}x^2}-1) - x\,e^{\frac{1}{2}x^2}\int_0^x e^{-\frac{1}{2}t^2}\,dt,$$

which is non-positive since $$e^{\frac{1}{2}x^2}\int_0^x e^{-\frac{1}{2}t^2}dt = \sum_{n=1}^{\infty} \frac{x^{2n-1}}{1.3.....(2n-1)}. \tag{5}$$

As a consequence, we obtain that this ratio (of the L.H.S. of (4) to the R.H.S) has an upper bound $2/\pi^{\frac{1}{2}}$.

2. A different lower bound for v can be obtained easily from a result proved by Chu & Hotelling (1954). There we showed that for all $x \geqslant 0$,

$$x^2(1-4v^2)/(4v^2) \leqslant \tfrac{1}{2}\pi. \tag{6}$$

Hence it follows that $$v \geqslant \tfrac{1}{2}[2x^2/(\pi+2x^2)]^{\frac{1}{2}}. \tag{7}$$

For easy reference, we will give here a proof of (6). Let

$$g_0(x) = x^2(1-4v^2)/(4v^2), \tag{8}$$

then $\lim_{x\to 0} g_0(x) = \frac{1}{2}\pi$. We will show that $g_0(x)$ is decreasing. Let a prime denote differentiation with respect to x. Then

$$g_0'(x) = x/(2v^3)\,g_1(x), \tag{9}$$

where $$g_1(x) = v(1-4v^2) - xv', \tag{10}$$

* Work sponsored by the Office of Naval Research under Contract NR 042 031 at Chapel Hill.

$$\left.\begin{aligned} &g_1(x) = g_2(x)\,v', \\ \text{where}\quad &g_2(x) = x^2 - 12v^2, \\ &g_2'(x) = (12/\pi)\,e^{-x^2} g_3(x), \\ \text{where}\quad &g_3(x) = (\pi/6)\,x\,e^{x^2} - e^{\frac{1}{2}x^2}\int_0^x e^{-\frac{1}{2}t^2}\,dt. \end{aligned}\right\} \tag{11}$$

From (5), we have $$g_3(x) = \sum_{n=0}^{\infty}\left\{\frac{\pi}{6n!} - \frac{1}{1.3.\ldots.(2n+1)}\right\} x^{2n+1}.$$

It can be shown, by a similar argument used by Pólya for a similar purpose, that

$$g_3(x) = x^3(a_0 x^{-2} + a_1 + a_2 x^2 + \ldots), \tag{12}$$

where $a_0 < 0$ and $a_i > 0$, $i = 1, 2, \ldots$. Hence there exists an $x_0 > 0$ such that $g_3(x) \leqslant 0$ if $0 \leqslant x \leqslant x_0$ and $g_3(x) \geqslant 0$ if $x \geqslant x_0$. So as x increases from 0 to ∞, $g_2(x)$ decreases steadily from 0 to a minimum and then increases steadily to ∞. Consequently $g_1(x)$ first decreases steadily and then increases steadily. As $\lim_{x\to 0} g_1(x) = \lim_{x\to\infty} g_1(x) = 0$, it becomes clear that $g_1(x) \leqslant 0$ for all $x \geqslant 0$. Therefore $g_0(x)$ is a decreasing function of x. Hence we have (6).

Comparison can be made easily of the two lower bounds for v given by (4) and (7). For simplicity, they will be denoted by $a(x)$ and $b(x)$ respectively. Now $a(x) \gtreqless b(x)$ according as $c(x) = e^{\frac{1}{2}x^2} - 2x^2/\pi - 1 \gtreqless 0$. As x varies from 0 to ∞, $c'(x)$, the derivative of $c(x)$, changes sign from negative to positive. So does $c(x)$. If $x = x_0$ is the solution of $c(x) = 0$, then $x_0 = 1$ approximately (the exact value is slightly smaller). Therefore, the lower bound in (7) is closer to v than that in (4) if $0 \leqslant x \leqslant 1$ (approximately) and less close if $x \geqslant 1$.

Further, the following statement is of similar nature to the one made in § 1.

If for all $x \geqslant 0$, $$v \geqslant \tfrac{1}{2}[ax^2/(1+ax^2)]^{\frac{1}{2}}, \tag{13}$$

then it is necessary and sufficient that $0 \leqslant a \leqslant 2/\pi$. On the other hand, for no finite a can the R.H.S. of (13) be, for all $x \geqslant 0$, an upper bound for v.

The above statement can be shown easily by considering the limit, as $x \to 0$, of the ratio of v to the R.H.S. of (13); and the limit, as $x \to \infty$, of $(1 - 4v^2)(1 + ax^2)$.

3. Several authors have derived inequalities for Mills's ratio. Their results can be written in the form of bounds for the normal integral. For example, in our notation, Gordon's (1941) inequalities are equivalent to

$$\frac{1}{2} - \frac{1}{x}(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \leqslant v \leqslant \frac{1}{2} - \frac{x}{x^2+1}(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \quad \text{for} \quad x > 0. \tag{14}$$

Birnbaum (1942) improved Gordon's upper bound in (14) and obtained

$$v \leqslant \frac{1}{2} - \frac{(4+x^2)^{\frac{1}{2}} - x}{2}(2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} \quad \text{for} \quad x \geqslant 0. \tag{15}$$

More recently, Tate (1953) showed what amounts to

$$\left(\frac{1}{4} + \frac{e^{-x^2}}{2\pi x^2}\right)^{\frac{1}{2}} - \frac{e^{-\frac{1}{2}x^2}}{x(2\pi)^{\frac{1}{2}}} \leqslant v \leqslant \tfrac{1}{2}(1 - e^{-x^2}) \quad \text{for} \quad x \geqslant 0. \tag{16}$$

We will now compare briefly (2) and (4) with (14), (15) and (16). The upper bound in (16) is obviously not so good as that in (2). The lower bound in (16) is non-negative for all real x. It is $\gtreqless$ the R.H.S. of (4) according as $h(x) = x^2 - (8/\pi)(1 - e^{-\frac{1}{2}x^2}) \gtreqless 0$, as will be seen by squaring the difference twice. As x varies from 0 to ∞, $h(x)$ decreases steadily from 0 to a minimum and then increases steadily to ∞; and vanishes at $x = 1{\cdot}01$ approximately (the exact value is slightly smaller). Therefore the lower bound in (16) is closer to v than that in (4) if and only if $x \geqslant 1{\cdot}01$ approximately.

The lower bound in (14) is an increasing function of x for all $x > 0$. It is non-negative when $x \geqslant 0{\cdot}65$ (approximately); and in this case it is $\gtreqless$ the R.H.S. of (4) according as $g(x) = 2(2/\pi)^{\frac{1}{2}} x - x^2 - (2/\pi) e^{-\frac{1}{2}x^2} \gtreqless 0$. As x varies from 0 to ∞, $g(x)$ increases steadily from $-2/\pi$ to a maximum and then decreases steadily to $-\infty$. The two roots of $g(x) = 0$ are approximately $x = 0{\cdot}5$ and $x = 1{\cdot}45$. Hence the lower bound in (14) is closer to v than that of (4) if and only if $x \geqslant 1{\cdot}45$ approximately.

Finally we point out that, for values of x close to 0, the upper bound in (2) is better than those in (14) and (15); while for large values of x, the latter two are better. No detailed comparison is attempted.

The author wishes to thank the referee for calling his attention to R. F. Tate's work and suggesting adding to the original note some comparison of the new and known results. Thanks are also due to Professor Harold Hotelling for his critical reading of the manuscript.

REFERENCES

BIRNBAUM, Z. W. (1942). *Ann. Math. Statist.* **13**, 245–6.
CHU, J. T. & HOTELLING, H. (1954). The moments of the sample median. To be published.
GORDON, R. D. (1941). *Ann. Math. Statist.* **12**, 364–6.
PÓLYA, G. (1949). *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, pp. 63–78. Berkeley: University of California Press.
TATE, R. F. (1953). *Ann. Math. Statist.* **24**, 132–4.
WILLIAMS, J. D. (1946). *Ann. Math. Statist.* **17**, 363–5.

Substitutes for χ^2

BY J. B. S. HALDANE

Department of Biometry, University College, London

Neyman (1930) and Jeffreys (1948, p. 170) have suggested a substitute for χ^2 involving some saving of computation. I here suggest what I believe to be a better one. If a sample consists of N individuals belonging to m classes, and n_r belong to the rth class, the expected number on some hypothesis being Na_r, where $\sum_{r=1}^{m} a_r = 1$, then

$$\chi^2 = \sum_{r=1}^{m} \frac{(n_r - Na_r)^2}{Na_r}.$$

Neyman's

$$\chi'^2 = \sum_{r=1}^{m} \frac{(n_r - Na_r)^2}{n_r}.$$

I consider

$$\chi''^2 = \sum_{r=1}^{m} \frac{(n_r - Na_r)^2}{n_r + 2}. \tag{1}$$

Since there is a finite probability that any n_r should be zero, it is clear that the expectation of χ'^2 is formally infinite. I shall show that it still exceeds $m-1$ even when samples in which any $n_r = 0$ are excluded. Haldane (1953) gave reasons for preferring $n_r + 1$ as a divisor in a similar context. It can be shown that

$$\mathscr{E}\left[\sum_r \frac{(n_r - Na_r)^2 + b}{n_r + c}\right] = m - 1 + N^{-1}[(b - c + 2)\Sigma a_r^{-1} - (3 - c)m + 1] + O(N^{-2}).$$

Hence to avoid an infinite expectation c must be positive, and to avoid a multiple of Σa_r^{-1}, which may be large, in the expectation, we must have $b = c - 2$. The value $b = 0$ gives a simple formula, though $b = 1$ gives an expectation nearer to $\mathscr{E}(\chi^2)$ when N is large.

Let $n_r = Na_r + x_r$. Then

$$\chi^2 = N^{-1}\sum_r x_r^2 a_r^{-1},$$

$$\chi'^2 = N^{-1}\sum_r x_r^2 a_r^{-1}\left(1 + \frac{x_r}{Na_r}\right)^{-1}$$

$$= \chi^2 + \sum_{i=2}^{\infty} [N^{-i}\sum_r (-x_r)^{i-1} a_r^{-i}],$$

$$\chi''^2 = N^{-1}\sum_r x_r^2 a_r^{-1}\left(1 + \frac{x_r + 2}{Na_r}\right)^{-1}$$

$$= \chi^2 + \sum_{i=1}^{\infty} [N^{-i-1}\sum_r x_r^2(-x_r - 2)^i a_r^{-i-1}].$$

To find the expectations of these quantities we require the expectations of powers of x_r, namely,

$$\mathscr{E}(x_r) = 0, \quad \mathscr{E}(x_r^2) = Na_r(1-a_r), \quad \mathscr{E}(x_r^3) = Na_r(1-a_r)(1-2a_r), \quad \mathscr{E}(x_r^4) = 3N^2a_r^2(1-a_r)^2 + O(N).$$

If we write $\mathscr{E}^*(x_r^i)$ to mean the expected value of x_r^i when n_r is not zero, we omit the cases where $x_r = -Na_r$, which have a probability $(1-a_r)^N$, which tends to zero quicker than any negative power of N. Thus

$$\mathscr{E}^*(x_r) = \frac{Na_r(1-a_r)^N}{1-(1-a_r)^N}, \quad \mathscr{E}^*(x_r^2) = \frac{Na_r(1-a_r)\,1-N^2a_r^2(1-a_r)^N}{1-(1-a_r)^N}, \quad \text{etc.}$$

So

$$\left.\begin{aligned} \mathscr{E}(\chi^2) &= N^{-1}\sum_{r=1}^{m}(1-a_r) = m-1, \\ \mathscr{E}(\chi'^2) &= \infty, \\ \mathscr{E}^*(\chi'^2) &= m-1+N^{-1}(2\Sigma a_r^{-1}-3m+1)+O(N^{-2}), \\ \mathscr{E}(\chi''^2) &= (m-1)\left(1-\frac{1}{N}\right)+O(N^{-2}). \end{aligned}\right\} \tag{2}$$

Thus even if we exclude the samples where any n_r is zero, χ'^2 has a positive bias often exceeding twice the reciprocal of the smallest expectation. The bias of χ''^2 is smaller, and readily calculated. The higher moments of the distribution of χ''^2 and of χ'^2, provided samples where any $n_r = 0$ are excluded, differ from those of χ^2 by quantities of the order N^{-1}. Errors of this order are neglected in the ordinary use of χ^2, and can be neglected in that of χ''^2, since χ^2 would be used if great precision were required.

As a numerical example, suppose that the numbers expected in four classes are 63, 21, 21 and 7, those observed being 71, 13, 16 and 12. Then $\chi^2 = 8{\cdot}825$, $\chi'^2 = 9{\cdot}470$, $\chi''^2 = 8{\cdot}319$. If we reverse the signs of the deviations, so that the observed numbers are 55, 29, 26 and 2, we find $\chi^2 = 8{\cdot}825$, $\chi'^2 = 16{\cdot}832$, $\chi''^2 = 10{\cdot}330$. The addition of the bias $0{\cdot}0268$ to χ''^2 gives values of $8{\cdot}345$ and $10{\cdot}357$, and this correction is clearly negligible. It is clear that χ''^2 is a far better approximation than χ'^2, and as it is no harder to calculate, it should be preferred.

REFERENCES

HALDANE, J. B. S. (1953). A class of efficient estimates of a parameter. *Bull. Int. Statist. Inst.* **33**, 231–48.

JEFFREYS, H. (1948). *Theory of Probability*. Oxford University Press.

NEYMAN, J. (1930). *Bull. Int. Statist Inst.* **24**, 44–86.

A problem in the significance of small numbers

BY J. B. S. HALDANE

Department of Biometry, University College, London

Dr H. Grüneberg and Dr A. G. Searle have each presented me with the following problem. N mice were examined anatomically. Before examination they had been classified into m groups, containing $n_1, n_2, n_3, \ldots, n_r, \ldots, n_m$ members. The members of a group were in fact grouped together on the basis of common ancestry. They could have been grouped on the basis of phenotypic resemblance. In one and only one of these groups, containing n members, a mice were found with a specific anatomical abnormality. What is the probability that such a coincidence should occur as the result of random sampling?

This question, like all such questions, is rather vaguely stated; and the answer to it may depend on the theory of probability adopted by the answerer. The answer here given is therefore not the only possible one.

We can arrange a fourfold table as follows:

	Normal	Abnormal
One group	$n-a$	a
Other groups	$N-n$	0

The application of Fisher's (1935) 'exact' method gives

$$P_1 = \frac{(N-a)!\,n!}{N!\,(n-a)!}$$

as the probability that all the abnormals should belong to this particular group. If they had all belonged to any one group this would have attracted notice provided P_1 was fairly small. If, however, one group had consisted of about half the total, that is to say $n = \frac{1}{2}N$ approximately, and a had not exceeded 3, the question of significance would not have arisen. The question can therefore be put somewhat more concretely: 'What is the probability that if there are a or more abnormals, all should occur in a group with n or less members?' Or still more definitely: 'If there were just a abnormals, and each group consisted of n members, what is the probability that by an accident of sampling, all the abnormals should be found in one group?'

The probability that all the abnormals should be found in one group is

$$P_2 = \frac{(N-a)!}{N!} \sum_{r=1}^{m} \frac{n_r!}{(n_r-a)!}.$$

If each group consists of just n members, which implies that N/n is an integer, this probability is

$$P_3 = P_1 \times \frac{N}{n} = \frac{N}{n} \frac{(N-a)!\,n!}{N!\,(n-a)!} = \frac{(N-a)!\,(n-1)!}{(N-1)!\,(n-a)!}.$$

I suggest that this is a reasonable estimate of the probability of the observed event, or of one equally or more unlikely, even if the values of n_r are unequal. We see that if $a = 1$, $P_3 = 1$, which is reasonable, since it is certain that the one abnormal will fall into one of the groups. We can then frame the question as follows: 'The first abnormal individual was found in a certain group. What is the probability that the $a-1$ abnormals found among the other $N-1$ mice should also be members of this group?' Clearly the value found, i.e. P_3, is a reasonable answer to this question.

In the case propounded to me by Dr Searle, $N = 472$, $n = 23$, $a = 2$, $m = 4$. So

$$P_3 = \frac{470!\,22!}{471!\,21!} = \frac{22}{471} = 0{\cdot}047.$$

The uncorrected value of P is $P_1 = 0{\cdot}0023$. At least one of the other groups must have consisted of 150 or more mice. Had the two abnormals belonged to this group no question of significance would have arisen. The estimation of P as about 0·05 rather than about 0·002 allows for the fact that, if the coincidence in question is an effect of random sampling, a number of other comparable coincidences would not be considered significant.

Any test of significance is somewhat arbitrary. For example, in place of the χ^2 test where

$$\chi^2 = \sum_{r=1}^{n} \frac{[a_r - \mathscr{E}(a_r)]^2}{\mathscr{E}(a_r)},$$

where a_r is an observed number and $\mathscr{E}(a_r)$ its expectation, we could use

$$\sum_{r=1}^{n} \frac{|a_r - \mathscr{E}(a_r)|}{[\mathscr{E}(a_r)]^{\frac{1}{2}}},$$

or many other similar statistics. χ^2 is used because it has a fairly simple distribution. On the same data the alternative might give a higher or a lower value of P. The test here suggested is easy to apply, and answers a reasonable question.

I think, however, that the problem merits a fuller discussion, and that a solution based on a different approach might be of equal or greater value.

REFERENCE

FISHER, R. A. (1935). The logic of inductive inference. *J. R. Statist. Soc.* **98**, 39–54.

Bounds for the ratio of range to standard deviation

BY GEORGE W. THOMSON

Ethyl Corporation, Detroit 20, Michigan

1. This note supplements the work of David, Hartley & Pearson (1954) on the distribution of the ratio of the range w to the standard deviation estimate s, where both are from the *same sample* of size n. Bounds are shown to exist for w/s for all populations with non-zero variance and percentage points are given for samples of three from a normal population.

2. The bounded nature of the distribution of w/s has not been noted by any of the authors who have investigated this statistic*. It can be readily shown that the upper and lower bounds for w/s for samples from *any* population with non-zero variance arise from certain simple configurations of sample points. The upper bound, which corresponds to minimum s for a given range w, results from the arrangement with $n-2$ of the points at the sample mean and the other two points at equal distances from the mean. The lower bound, which corresponds to maximum s for a given w, results from the concentration of half

Table 1. *Bounds of the distribution of the ratio of range to standard deviation, w/s, in samples of size n from a normal population*

n	Percentage points		n	Percentage points		n	Percentage points	
	Lower 0%	Upper 0%		Lower 0%	Upper 0%		Lower 0%	Upper 0%
			10	1·897	4·243	30	1·966	7·616
			11	1·915	4·472	40	1·975	8·832
			12	1·915	4·690	50	1·980	9·899
3	1·732	2·000	13	1·927	4·899	60	1·983	10·863
4	1·732	2·449	14	1·927	5·099	80	1·987	12·570
						100	1·990	14·071
5	1·826	2·828	15	1·936	5·292			
6	1·826	3·162	16	1·936	5·477	150	1·993	17·263
7	1·871	3·464	17	1·944	5·657	200	1·995	19·950
8	1·871	3·742	18	1·944	5·831	500	1·998	31·591
9	1·897	4·000	19	1·949	6·000	1000	1·999	44·699
			20	1·949	6·164			

of the sample points at one extreme and the other half (plus one, if the sample size is odd) of the sample points at the other extreme. The numerical values of the bounds can be shown to be:

$$\text{Upper bound of } w/s: \quad \sqrt{\{2(n-1)\}}.$$

$$\text{Lower bound of } w/s: \quad 2\sqrt{\{(n-1)/n\}} \quad \text{for } n \text{ even},$$

$$2\sqrt{\{n/(n+1)\}} \quad \text{for } n \text{ odd}.$$

As n becomes larger the lower bound approaches 2. It is also evident that these bounds are distribution-free provided that the points can be distributed at all. Table 1 shows the numerical values which correspond to the lower and upper 0 % points in Table 6 of the paper by David *et al.*

* [*Editorial Note.* The existence of these limits has no doubt been noticed by others who have considered the distribution of this ratio; in the correspondence leading to the joint paper by David, Hartley & Pearson (1954), the first author gave these limits in a letter of February 1954, but they were omitted in the published paper. E.S.P.]

3. Certain properties of the distribution of samples of size 3 from a normal population have been obtained by Lieblein (1952), as noted by Seth (1950). It is easily shown from Lieblein's results that the percentage points of w/s for $n = 3$ are given by

$$2\cos[30°(1-F)],$$

where F is the cumulative frequency. Thus, for the upper 10 % point, $F = 0{\cdot}90$, $w/s = 2\cos 3° = 1{\cdot}99726$. Table 2 shows the upper and lower 0·0, 0·5, 1·0, 2·5, 5·0 and 10·0 % points and the median. The upper percentage points agree to the third decimal place with those given by David *et al.* (1954, Table 6).

Table 2. *Percentage points of the distribution of the ratio of range to standard deviation, w/s, in samples of size 3 from a normal population*

Lower percentage points	w/s	Upper percentage points	w/s
0·0	1·73205	10·0	1·99726
0·5	1·73466	5·0	1·99931
1·0	1·73726	2·5	1·99983
2·5	1·74499	1·0	1·99997
5·0	1·75763	0·5	1·99999
10·0	1·78201	0·0	2·00000
Median			
50·0	1·93185		

4. The bounds for w/s have been used in our research laboratory for the past twelve years in routine checks of the computation of s for small samples, since gross errors can be detected at once.

REFERENCES

David, H. A., Hartley, H. O. & Pearson, E. S. (1954). *Biometrika*, **41**, 482–93.
Lieblein, J. (1952). *J. Res. Nat. Bur. Stand.* **48**, 255–68.
Seth, G. R. (1950). *Ann. Math. Statist.* **21**, 298–301.

On the estimation of population parameters from marked members

By J. A. GULLAND

Fisheries Laboratory, Lowestoft

In commercial fish populations there exist two distinct types of mortality, that due directly to capture by man, and that due to other causes. For simplification it is not unreasonable to assume that these may be represented by the constant exponential coefficients, denoted by F and M. It is the separation and estimation of these which forms a major part of any study of such populations.

Typically, the total population is of the order of millions, and recaptures from a marking experiment consist of all the marked fish caught by any fisherman and recaptures may therefore be considered as a continuous process rather than at discrete intervals, as in most experimental trapping (e.g. Hammersley, 1953). If a known number of marked individuals is released, then direct maximum likelihood estimates for the two mortalities F and M (in the marked population) may be obtained.

For the probability of an individual remaining alive until time t, when released at time $t = 0$, is $e^{-(F+M)t}$. Hence the probability of recapture in interval $(t, t+dt)$ is $F\,e^{-(F+M)t}\,dt$. Suppose now N individuals are released, of which n are recaptured at time $t_1, t_2, \ldots, t_n$, and that these are the only recaptures up to time T ($T >$ all t), when the experiment was ended. Then the probability of being recaptured is

$$\int_0^T F\,e^{-(F+M)t}\,dt = \frac{F}{F+M}(1-e^{-(F+M)T}),$$

and the probability of not being recaptured is

$$\frac{M}{F+M}+\frac{F}{F+M}e^{-(F+M)T}.$$

The likelihood function may therefore be written

$$e^{L}=\binom{N}{n}\prod_{i=1}^{n}(F\,e^{-(F+M)t_i})\left(\frac{M}{F+M}+\frac{F}{F+M}e^{-(F+M)T}\right)^{N-n}.$$

Hence $$L=\log\binom{N}{n}+n\log F-(F+M)\,\Sigma t_i+(N-n)\{\log(M+F\,e^{-(F+M)T})-\log(F+M)\},$$

$$\frac{\partial L}{\partial F}=\frac{n}{F}-\Sigma t_i+(N-n)\left\{\frac{(1-FT)\,e^{-(F+M)T}}{M+F\,e^{-(F+M)T}}-\frac{1}{F+M}\right\}$$

and

$$\frac{\partial L}{\partial M}=-\Sigma t_i+(N-n)\left\{\frac{1-FT\,e^{-(F+M)T}}{M+F\,e^{-(F+M)T}}-\frac{1}{F+M}\right\}.$$

If the experiment is continued so long that $e^{-(F+M)T}$ may be neglected, then these equations become very simple, namely,

$$\frac{\partial L}{\partial F}=\frac{n}{F}-\Sigma t_i-\frac{N-n}{F+M},\quad \frac{\partial L}{\partial M}=-\Sigma t_i+\frac{(N-n)\,F}{(F+M)\,M}.$$

Putting these equal to zero gives the solution

$$\hat{F}=\frac{n^2}{N\Sigma t_i},\quad \hat{M}=\frac{(N-n)\,n}{N\Sigma t_i}.$$

These estimates are biased, in fact having infinite expectations and higher moments. However, $\Sigma t_i/n$ is an unbiased estimate of $(F+M)^{-1}$, when $e^{-(F+M)T}$ is neglected, and in practice, for reasonably large values of n and N the distribution of $\hat{F}$ and $\hat{M}$ for repeated experiments is likely to be quite reasonable.

REFERENCE

HAMMERSLEY, J. M. (1953). *Biometrika*, **40**, 265–78.

REVIEWS

Biomathematics. By CEDRIC A. B. SMITH. London: Charles Griffin and Co. Ltd. Pp. xv+712. 80*s*. net.

Biomathematics, by the late Dr W. M. Feldman, appeared in 1923 and reached a second edition in 1935. The present book is modestly described by Dr Smith as a third edition. But it has been entirely rewritten and is, in effect, a new book.

Although advanced statistical theory uses nearly all branches of mathematics, except perhaps projective and differential geometry, it draws on them to very varying extents. The student of statistics can go a long way with only a nodding acquaintance with analytic geometry and differential equations, but finds a knowledge of calculus and advanced algebra almost essential from the outset. He therefore requires a special course of mathematics for statisticians and will not find it within the covers of any single book. There is a great and growing need for a work of this kind.

Dr Smith's book does for biologists what one would like to see done for statisticians in general. It provides a thorough grounding in the ideas and techniques necessary for those who want to be mathematical biologists without going so far as to be biological mathematicians; and it does so with great fluency and insight without any sacrifice of rigour. There can be few writers as well qualified as Dr Smith to write such a work. His mathematical powers, his extensive knowledge and practical experience of mathematical and statistical applications in biology, the care he has bestowed on the work and his gifts as a teacher combine to make this a work of outstanding excellence. It will be useful not only to biologists but to any student who requires a competence in mathematics for the purposes of pursuing his own subject.

The book opens with two chapters on arithmetic, including some calculating devices and some account of mechanical equipment and punched card machines. Chapter 3 is a refresher course on algebra, leading to the treatment of inequalities in Chapter 4. Chapter 5 deals with the connexion between algebra and geometry and provides a foundation in analytic geometry and conic sections. The next two chapters deal with logarithms (which Dr Smith introduces by means of the equi-angular spiral) and their applications in computation by slide-rules and nomograms. Chapters 8–12 are on differential and integral calculus and proceed as far as simple differential equations and partial differentials. Chapter 13 deals with series and Chapter 14 with vectors. Methods of solving equations and matrices receive separate chapters, and the book closes with three chapters on Chance, Statistical Distributions and Simple Statistical Procedures, together with one on Colson's method of simplifying arithmetical calculations. The book contains over 700 pages and the amount of ground covered is astonishing considering that no impression is given of hurry or over-consideration.

Readers of Stephen Leacock will remember his protests about the dullness of arithmetical examples, and his attempt to dispel it by investing with human attributes those three anonymous characters of our youth *A*, *B* and *C*; especially the case of poor *C*, who died of pneumonia contracted while trying to fill a cistern with a leak in it. Dr Smith, who combines with his other qualities a sense of humour, would have endeared himself greatly to Leacock. For instance, the additive properties of matrices are illustrated on the population of Narkover, the determination of minima is exemplified, not on those eternal salmon tins, but on the angles of origin of human arteries; the treatment of physical magnitudes is enlivened by an example leading to the conclusion that if one falls into a cold sea the best thing to do is to swim hard so as to keep warm.

Altogether, this is a very sound, readable and sensible book. In spite of its price it should become widely popular.

M. G. KENDALL

Introduction to Mathematical Statistics. 2nd edition. By PAUL G. HOEL. New York: John Wiley and Sons Inc.; London: Chapman and Hall. 1954. Pp. ix+331. 40*s*.

This is a much expanded and revised version of a book that first appeared in 1947. Apart from some expansion of many topics such as regression, the analysis of variance and non-parametric methods, there have been notable additions in the probability field. As a result the book is much more self-contained than formerly. It presupposes a knowledge of calculus—probably about two years study—and seems designed for scientists who have some mathematical background and wish to master the

elements of statistical theory. For the applications to their own fields, they will have to turn to other texts.

Inevitably in a book of this size some things have had to be omitted, for example, time series, equalization of variance and probit analysis. It is perhaps surprising, in view of the intended public, that no mention is made of Sheppard's corrections for grouping or the many short-cut procedures available nowadays, such as the range in the analysis of variance. There is also an unusual differentiation between large sample and small sample distributions—is the F distribution a small sample distribution? Diagrams are used excellently throughout the book and are well labelled. There are numerous exercises for which no answers are provided. The book ends with a number of useful tables, but whilst the percentage points of t, F and χ^2 are given, those for the normal curve have to be obtained by inverse interpolation in the table of the cumulative distribution provided.

P. G. MOORE

Design and Analysis of Industrial Experiments. Edited by O. L. DAVIES. London and Edinburgh: Oliver and Boyd, for Imperial Chemical Industries Limited. 1954. Pp. xiii + 636. 63*s*.

This book is written by a team of authors largely the same as those who wrote *Statistical Methods in Research and Production* in 1947, but, as explained in the Introduction, the present volume deals with the design of experiments and their subsequent analysis rather than the extraction of information from previously existing data. The particular point of the book is that it is written from a broadly chemical point of view, rather than the more usual agricultural one, and so the design of experiments is freed from the restrictions imposed by agricultural conditions, and sometimes inadvertently carried over into fields where they do not apply.

Successive chapters deal with Simple Comparisons, Sequential Tests—a remarkably fine chapter—Sampling and Testing Methods, Randomized Blocks, and Incomplete Randomized Blocks, again very good because it concentrates attention on the simpler designs required for industrial experimentation. It is nice to have the assurance that these are worth while, for the alternative view that in them too few observations have to bear a heavy load of theoretical interpretation is somewhat prevalent. Then there are four chapters on Factorial Experiments and one on the determination of optimum conditions. In this last chapter the treatment is largely orographical, which I like as it enables simple words like 'ridge' to be used as short-cuts avoiding much difficult mathematical explanation. Then follows a glossary which is far the worst part of the book. No readers of a book of this type need to have explained what is quaintly called 'The arithmetic average or mean', and if they do the definition here given will hardly help them, nor is it possible to explain 'Universe, Population, Parameter, Sample, Statistic' in fourteen lines. The final tables are well arranged.

The book as a whole contains a very large amount of valuable information, but as might be expected from a team the style varies greatly in clearness and readability; certainly the best parts are very good. There is an implied assumption that nothing but the experiment and its own conditions need be considered in planning, and this is certainly not always true for those who work in less magnificent concerns than I.C.I., but still you must know how the experiment really ought to be done, before planning a compromise with what can be done, and the necessary information is here.

Finally, it must be emphasized that in industry an investigation is not complete when the results have been obtained in conventional statistical form. It is necessary to translate them into the language used by the executive who is to decide what action to take on them. It is avowedly not the purpose of this book to deal with this question, but there is internal evidence that at least one of the authors would be well able to do so, and I should like to suggest that in a future edition such a chapter should be added in place of the glossary.

L. MCMULLEN

Sample Survey Methods and Theory. By M. H. HANSEN, W. N. HURWITZ and W. G. MADOW. New York: John Wiley and Sons Inc.; London: Chapman and Hall. 1953. Vol. I, **Methods and Applications.** Pp. xxii + 638. 64*s*. Vol. II, **Theory.** Pp. xiii + 332. 56*s*.

These books aim to cover the whole field of the sample survey and cater for all tastes and all classes of statistician. Vol. I sets out the methods in common use in sampling surveys and the way in which the methods are applied. It begins with commonsense talk about the fundamental aims of sampling and sampling design and continues with the delineation of such statistical ideas as may be expected to be

useful to the particular purpose of the authors. The chapter on bias and non-sampling errors in survey results is obviously written from practical experience. Much of the book is concerned with descriptions of the various types of sampling—simple random sampling, stratified simple random sampling, simple one or two stage cluster sampling, stratified single or multi-stage cluster sampling. In each case the procedure is clear, every possible query which might be raised by the would-be user is answered. The chapter on estimating variances and the accuracy of the method of estimation will not be helpful unless the reader has learned a little elementary statistics but presumably if he has reached this point in the book he will willy-nilly have acquired such basic information as he will here find to be necessary.

Vol. II can either be read in conjunction with Vol. I or separately when it can be looked on as an elementary text-book exercise in statistical algebra. The standard here is uneven. It is surely unnecessary to explain 'why we study summation notation' (p. 11) to someone who is expected to understand without explanation why we study convergence in probability (p. 72). None the less in spite of irrelevancies the algebraic framework is clear and concise and should not cause undue difficulty to any student interested enough to read the book. Certain contractions are possibly invented by the writers and should be rejected; 'plim' standing for 'limit in the sense of probability' is one of the least useful of these. All the formulae used in vol. I are derived here, and the way is clear for the reader to make such modifications as his own particular sampling design may dictate.

The two books taken together cover fully the whole field of sample surveys, and are unlikely to be superseded for some years to come. This being the case it is unfortunate that such encyclopaedic works should be priced out of the possible range for many persons.

F. N. DAVID

Sampling Theory of Surveys with Applications. By PANDURANG V. SUKHATME. New Delhi: The Indian Society of Agricultural Statistics; Ames, Iowa: The Iowa State College Press. 1954. Pp. xxviii + 491. $6.

This is the fifth book dealing specifically with sample survey methods and theory to appear in the last five years. The author is head of the statistics branch of F.A.O. and has wide experience of sample surveys in India and elsewhere, and one might have expected his book to be more closely concerned than it is with the methods of overcoming the difficulties arising in agricultural surveys and in surveys of underdeveloped territories. In fact, Dr Sukhatme has written a text-book on the algebraical development of the standard branches of sample-survey theory. On this basis the treatment is painstaking and clear and the book will be most valuable to the student or research worker who requires an ad hoc treatment of the theory appropriate to any particular sampling method.

The up-to-date character of the book is shown by the prominence given to unequal-probability sampling, the treatment of which is based in part on the author's own research, and by the chapter on non-sampling errors, which contains some novel results on the treatment of interviewing errors.

However, it is to be regretted that Dr Sukhatme in common with other recent writers on sample survey theory has neglected the opportunity to base his treatment of the theory on the simplifying principles set out by Yates five years ago in *Sampling Method for Censuses and Surveys*. Writing mainly for practical workers rather than for theoreticians—no proofs of the formulae were given—Yates based his exposition on the fact that most survey designs are built up of the following components:

(*a*) Method of selection, e.g. simple random, systematic, probability proportional to size.
(*b*) Stratification.
(*c*) Use of supplementary information.
(*d*) Multi-stage sampling.
(*e*) Multi-phase sampling.

The number of combinations of the various possibilities under each of these heads is so large that it would be impossible to give within reasonable bounds the appropriate estimation and variance formulae for all sample designs likely to be useful in practice. However, Yates pointed out that the contribution from each component could be considered separately and the results combined to give the relevant formulae for any particular design.

As an example, the estimate of a mean based on a stratified sample normally takes the form $\bar{y} = \Sigma N_i \bar{y}_i / \Sigma N_i$, where $\bar{y}_i$ is the estimate of the ith stratum mean. Suppose that by referring to the theory of unstratified sampling under (*a*) above, we were able to calculate an estimate s_i^2 of $V(\bar{y}_i)$. Yates would immediately write down the estimate of $V(\bar{y})$ as $\Sigma N_i^2 s_i^2 / (\Sigma N_i)^2$, and would assert that this result held generally whatever methods are used for sampling within strata. Sukhatme, however, has preferred to start from the beginning for each type of stratified sample and to develop explicit formulae

from first principles for each case, rather than prove a general rule for all types of stratified sample. Similarly, in his treatment of both the ratio method and multi-stage sampling, the author has preferred to treat each case separately from first principles rather than to develop the general rules by which the estimates can be built up. This is all the more regrettable since the general laws are easier to prove than the formulae for special cases.

A further opportunity for simplification was missed in the treatment of sampling with unequal probabilities with replacement. If we draw a sample of n with replacement from a population of values $y_1, y_2, \ldots, y_N$ with selection probabilities $p_1, p_2, \ldots, p_N$, all necessary estimation and variance formulae can be obtained by regarding the sample as a simple random sample from the discrete probability distribution in which the observation y_i has associated with it the probability p_i. Thus, once the theory of sampling from an infinite population with equal probabilities of selection has been worked out the corresponding results for unequal-probability sampling with replacement emerge as corollaries. In Dr Sukhatme's treatment, however, these results are worked out afresh.

For these reasons the book contains a great deal of avoidable algebra. Although some students undoubtedly find it helpful to have an ad hoc theory set out for each particular type of sample design, the consequent details are likely to prove unnecessarily time-consuming for the reader whose object is to grasp the principles of sample-survey theory.

J. DURBIN

Life and other Contingencies, Vol. I. By P. F. HOOKER and L. H. LONGLEY-COOK. Cambridge University Press. 1953. Pp. viii + 312. 22*s*. 6*d*.

This book is one of a series commissioned by the Institute of Actuaries and the Faculty of Actuaries to provide a course of reading suitable for the examinations conducted by these bodies. The present work replaces Spurgeon's *Life Contingencies*, which now ends an honourable career of some forty years. The authors have evidently been influenced to a considerable extent by their distinguished predecessor. The main changes arise from the demands of the new actuarial syllabus and consist of the introduction of sections on extra risks, valuation methods and non-mortality (sickness, maternity, etc.) benefits. There are also a number of additions necessitated by developments in theory and practice, such as the treatment of family income benefits and the appendix on International Actuarial Notation.

However, the book is by no means merely a rather thorough revision of the earlier text-book. It is a completely new work and develops the subject in its own way. Emphasis is definitely on the provision of methods for practical application in life-office work. There is little preoccupation with complicated algebraic manipulation, and the theoretical study of mortality laws and stationary populations is reduced to a minimum. Worked examples are collected at the ends of chapters and, perhaps for this reason, appear to be less profuse than in Spurgeon's text-book. There are no exercises to be worked out by the student.

The treatment of multiple decrement tables is left to vol. II, and it will be interesting to see how the authors tackle this still somewhat controversial topic. As a result of this division of subject-matter joint-life assurances and annuities are not considered in the present volume.

It is to be hoped that 'Hooker and Longley-Cook' will be as long-lived as 'Spurgeon'. If this is so the reviewer hopes to see in the successive revisions a much fuller treatment of non-mortality benefits, and, if possible, either an earlier and more adequate discussion of the construction of tables, or the omission of this subject by its transfer to another part of the actuarial syllabus.

N. L. JOHNSON

Population Statistics and their Compilation. By HUGH H. WOLFENDEN. The University of Chicago Press, for the Society of Actuaries. 1954. Pp. xxii + 258. 56*s*. 6*d*.

This book is a completely revised edition of a work originally published as 'Actuarial Study No. 3' by the Actuarial Society of America in 1925. The choice of subject-matter clearly reflects the actuarial interests of the author. In the analytical section of the book there is much greater emphasis on the study of mortality than on any other aspect of population phenomena. No fewer than 70 pages are devoted to the discussion of various methods of constructing mortality-tables and life-tables, while a further 23 pages are concerned with the comparison of mortality by occupations, causes, etc. and a brief discussion on forecasting mortality rates. There are only eight pages on the use of data on marriages, births, orphanhood & unemployment and three pages on 'Sickness Data'. Emigration and immigration are not specifically considered at all. The theory of reproductivity is relatively adequately

treated in 16 pages containing some sensible warnings on the uncritical use of the various indices of reproductivity.

The earlier part of the book shows much less evidence of this lack of balance. There is an interesting and useful description of the compilation of population statistics by means of censuses, and registration of births, marriages and deaths. This is accompanied by a discussion of likely sources of unreliability in the data, and methods of correcting the consequent defects.

Comprehensiveness has evidently been the author's ideal. While he has succeeded in making his book a mine of information on particular topics, it is to be expected that students may experience difficulty in sorting out really important material from items now mainly of historical interest. The dust-cover claims that 'this book is the only presentation, by an actuary, of the particular actuarial viewpoints and methods necessary to the production of modern population statistics'. Apart from the fact that this statement ignores books published by the Institute of Actuaries in recent years, the catalogue-like assembly of methods presented makes difficult the disentanglement of the more modern methods.

Despite these criticisms, the book should prove useful as a comprehensive, yet handy, work of reference. There is no index, though there is a detailed 'Table of Contents' by paragraphs (not by pages). There are nearly one hundred footnotes which bear witness to the author's devotion to the ideal of comprehensiveness. Most of them contain useful information, though they tend to interrupt the textual line of argument.

There is a 22-page appendix on 'Some Theory in the Sampling of Human Populations', by W. E. Deming. It is well written but will be of little value except to those with sufficient statistical training and such persons will probably be familiar with the subject already.

N. L. JOHNSON

Table of Binomial Coefficients. Royal Society Mathematical Tables, 3. London: Cambridge University Press. 1954. Pp. viii + 162. 35*s*.

These tables, prepared under the editorship of J. C. P. Miller, were computed mainly at the National Physical Laboratory, Liverpool University, and the Royal Aircraft Establishment at Farnborough. They give ${}^nC_r$ for

(i) $r \leqslant \frac{1}{2}n \leqslant 100$,
(ii) $2 \leqslant r \leqslant 12$, $200 \leqslant n \leqslant 500$,
(iii) $2 \leqslant r \leqslant 11$, $500 \leqslant n \leqslant 1000$,
(iv) $2 \leqslant r \leqslant 5$, $1000 \leqslant n \leqslant 2000$,
(v) $r = 2, 3$, $2000 \leqslant n \leqslant 5000$.

The computations throughout are performed by use of the recurrence relation

$$ {}^{n+1}C_r = {}^nC_r + {}^nC_{r-1}. $$

The main application of the table lies in the field of number theory for such investigations as to the possible uses of tetrahedral numbers to form other numbers. For statisticians the combinatorial numbers mainly conjure up visions of binomial probabilities and here, as the probabilities have to be combined with the combinatorial number, the volume is unlikely to supplant such well-worn friends as the incomplete beta function, at any rate for the lower values of n. Occasions do, however, arise where binomial coefficients alone are required—certain probability problems of a combinatorial nature for example—and these tables will form a useful reference work for such cases. The volume is well set out with very clear type, and a useful index is provided to aid one in locating the value required.

P. G. MOORE

Bessel Functions and Formulae. Compiled by W. G. Bickley. London: Cambridge University Press, for the Royal Society. 1953. Pp. 11. 3*s*. 6*d*.

This collection of formulae is a straight reprint of the *Summary of Notations* and the section on 'Functions and Formulae' from the British Association Mathematical Tables, Vol. x, Bessel Functions, ii.

It gives a summary of the differential, integral and difference relations obeyed by the Bessel and auxiliary functions and also some connecting formulae with related functions (e.g., Hankel's, Whittaker's, etc.). It further gives a variety of series expansions for and involving Bessel functions.

Of particular interest to statisticians are the sections giving generating functions and Laplace transforms of Bessel functions.

D. E. BARTON

PUBLICATIONS OF U.S. DEPARTMENT OF COMMERCE,
NATIONAL BUREAU OF STANDARDS

(i) Tables of Chebyshev Polynomials $S_n(x)$ and $C_n(x)$. Applied Mathematics Series, 9. 1952. Pp. xxix + 161. $1.75.

Defining the nth order polynomials as

$$C_n(x) = 2\cos\{\arccos \tfrac{1}{2}x\}, \quad S_n(x) = 2(4-x^2)^{-\frac{1}{2}}\sin\{(n+1)\arccos \tfrac{1}{2}x\},$$

the main tables gives their values to 12 decimal places for $m = 2(1)12$, $x = 0(0{\cdot}001)2$. Subsidiary tables give the expansions of these polynomials in powers of x, together with those of the modified forms:

$$T_n(x) = \tfrac{1}{2}C_n(2x), \quad T_n^*(x) = \tfrac{1}{2}C_n(4x-2),$$
$$U_n(x) = S_n(2x), \quad U_n^*(x) = S_n(4x-2).$$

Also given are inverse relations for powers of x as linear sums of $T_n(x)$ and $T_n^*(x)$ up to the twelfth power.

These polynomials are chiefly of use as ancillaries to the computing of functions possessing power series expansions over a whole range of values of the argument, as the convergence of a series of Chebyshev polynomials is more uniform and in general more rapid than that of a power series expansion taken about a point of the range. The computation was carried out under the direction of A. N. Lowan and a 22-page introduction is written by Cornelius Lanczos. This gives, *inter alia*, the example of the asymptotic expansion of the incomplete normal integral, where the Chebyshev series to six terms gives one-fifth the error of approximation of the power series expansion of the same order, for the possibility of deviating more than $\sqrt{2}$ standard deviations from the mean. These polynomials are distinct from those commonly used in statistical practice and also described as Chebyshev's, defined by

$$P_m(x) = \Delta^m\left[\binom{x+\frac{1}{2}(n-1)}{m}\binom{x-\frac{1}{2}(n+1)}{m}\right].$$

(ii) Tables of coefficients for the numerical calculation of Laplace transforms. Applied Mathematics Series, 30. 1953. Pp. 36. 25 cents.

These tables give coefficients for the approximate evaluation by quadrature of the Laplace transform

$$F(p) = \int_0^\infty f(t)\,e^{-pt}\,dt$$

of a function $f(t)$ which is given for n equally spaced values of t whose range includes the effective range of the argument. The coefficients are given to nine decimal places for $n = 2(1)11$ and varying ranges and intervals of p depending on n (e.g. $p = 0{\cdot}1(0{\cdot}1)1{\cdot}0$ for $n = 2$, $p = 1(1)10$ for $n = 11$). Special tables are also given for the simpler case where $f(t)$ is a low order polynomial. A short introduction is provided by H. E. Salzer. In this he remarks: 'Convenient estimates of the error of approximation seem difficult to obtain.' However, in the example given of $f(t) = J_0(t)$ it is found that for $n = 11$, $F(p)$ differs from the approximating function by less than 0·14 % over the range of p considered.

D. E. BARTON

iii) Tables of Lagrangian Coefficients for sexagesimal interpolation. Applied Mathematics Series, 35. 1954. Pp. ix + 157. $2.00.

These tables give 3-, 4-, 5- and 6-point Lagrangian interpolation coefficients A_t for arguments in sexagesimal measure, such as angles given in units of degrees, minutes and seconds. The coefficients are given for 3600 values of the fraction of the tabular interval. Thus if the function is tabled for each degree, the coefficients may be used to find a value of the function at any required minute and second. Each coefficient is tabled to eight decimal places. There is a brief Introduction by H. E. Salzer.

(iv) Tables of circular and hyperbolic sines and cosines for radian arguments. Applied Mathematics Series, 36. 1953. Pp. x+407. $3.00.

The main Table I gives for $x = 0(0{\cdot}0001)1{\cdot}9999$ values to nine decimal places of the functions $\sin x$, $\cos x$, $\sinh x$ and $\cosh x$. There are three short supplementary tables:

Table II gives values to nine decimals of the same four functions for $x = 0{\cdot}0(0{\cdot}1)10{\cdot}0$.
Table III is a conversion table for expressing degrees, minutes and seconds in radians and vice versa.
Table IV gives to 15 decimal places values of $n \times \frac{1}{2}\pi$ for $n = 1(1)100$.

There is a brief Introduction by A. N. Lowan.

(v) Table of secants and cosecants to nine significant figures at hundredths of a degree. Applied Mathematics Series, 40. 1954. Pp. vi+46. 35 cents.

This tables gives $\sec x$ and $\operatorname{cosec} x$ for $x = 0{\cdot}00(0{\cdot}01)90{\cdot}00$ degrees. It will serve as a companion to the table of $\sin x$ and $\cos x$ to fifteen decimal places at hundredths of a degree previously published as No. 5 in the Applied Mathematics Series.

E. S. PEARSON

CORRIGENDA

Biometrika (1954), **41**

(1) M. E. Wise, p. 328. For equation (8·5) read:

$$Nx = (N - \tfrac{1}{2}n + \tfrac{1}{2})(1-h) + \tfrac{1}{2}(p-1)$$

(2) P. Whittle, p. 437. On the left-hand side of equation (16)

for $\xi_t - a\xi_{t-1} - b\xi_{t-2}$ *read* $\xi_t - a\xi_{t-1} - b\xi_{t+1}$

(3) H. A. David, p. 466. In Table 2, the result for the rectangular population with $n = 4$

for 1·019 *read* 1·010

(4) D. R. Cox, p. 472. In Table 2, for population (*b*), the value for $n = 4$:

for 1·961 *read* 1·939, and for $n = 5$: *for* 2·252 *read* 2·196

The Corrigenda to papers by Rushton and Ruben stated in the List of Contents for Vol. **41**, Parts 3 and 4, as printed on p. 568, were wrongly placed *in front* of the first page (p. 287) of that issue.

(All Rights reserved)

BIOMETRIKA. Vol. 42, Parts 1 and 2

CONTENTS

Studies in the history of probability and statistics. I. Dicing and gaming (a note on the history of probability). By F. N. David 1

Some features of the generation times of individual bacteria. By E. O. Powell 16

Quantum hypotheses. By S. R. Broadbent 28

The truncated negative binomial distribution. By M. R. Sampford 58

The randomization analysis of a generalized randomized block design. By M. B. Wilk 70

Some quick sign tests for trend in location and dispersion. By D. R. Cox and A. Stuart 80

The variance of the maximum of partial sums of a finite number of independent normal variates. By A. Z. Anis 92

Spatial point processes, with applications to ecology. By H. R. Thompson 102

The outcome of a stochastic epidemic—a note on Bailey's paper. By P. Whittle 116

A note on Bailey's and Whittle's treatment of a general stochastic epidemic. By F. G. Foster 123

The deterministic model of a simple epidemic for more than one community. By S. Ruebush and A. J. Mautner 129

Exact tests for serial correlation. By E. J. Hannan 133

On the efficiency of procedures for smoothing periodograms from time series with continuous spectra. By M. S. Bartlett and J. Medhi 143

The autocorrelation function and the spectral density function. By J. Wise 151

Sampling properties of local statistics in stationary stochastic series. By G. H. Jowett 160

Models for two-dimensional stationary stochastic processes. By V. Heine 170

Some problems in the theory of provisioning and of dams. By J. Gani 179

Approximate confidence intervals. III. A bias correction. By M. S. Bartlett 241

The theory of correlation between two continuous variables when one is dichotomized. By Robert F. Tate 205

The normal probability function: tables of certain area-ordinate ratios and of their reciprocals. Editorial 215

Tables of symmetric functions. Part V. By F. N. David and M. G. Kendall 222

Control charts with warning lines. By E. S. Page 243

Miscellanea

Approximations to the probability integral and certain percentage points of a multivariate analogue of Student's t-distribution. By C. W. Dunnett and M. Sobel 259

Galton's rank-order test. By J. L. Hodges, Jr. 261

On bounds for the normal intergral. By John T. Chu 265

Substitutes for χ^2. By J. B. S. Haldane 265

A problem in the significance of small numbers. By J. B. S. Haldane 268

Bounds for the ratio of range to standard deviation. By George W. Thomson 268

On the estimation of population parameters from marked members. By J. A. Gulland 269

Reviews

Cedric A. B. Smith's 'Biomathematics' 271

Paul G. Hoel's 'Introduction to Mathematical Statistics' 271

O. L. Davies (ed.), 'Design and Analysis of Industrial Experiments' 272

M. H. Hansen, W. N. Hurwitz and W. G. Madow's 'Sample Survey Methods and Theory' 272

Pandurang V. Sukhatme's 'Sampling Theory of Surveys with Applications' 272

P. F. Hooker and L. H. Longley-Cook's 'Life and other Contingencies' 273

Hugh H. Wolfenden's 'Population Statistics and their Compilation' 273

Royal Society, 'Table of Binomial Coefficients' 274

W. G. Bickley's 'Bessel Functions and Formulae' (for the Royal Society) 274

National Bureau of Standards, Publications of the U.S. Department of Commerce, Applied Mathematics Series 9, 30, 35, 36, 40 275

Corrigenda 276

Printed in Great Britain at the University Press, Cambridge (Brooke Crutchley, *University Printer*)

REPRINTED IN BELGIUM BY JOS.ADAM